11-4-2019

Subjective Image Quality Assessment of Digitally Printed Images

Gaurav Sheth
gds6536@rit.edu

R.I.T

Subjective Image Quality Assessment of Digitally Printed Images

by

Gaurav Sheth

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Science
in Color Science

Program of Color Science
College of Science
Rochester Institute of Technology
Rochester, NY

Date of Completion 11/04/2019

Accepted by ______________________________
Dr. Mark Fairchild, Graduate Program Director Date

COLLEGE OF SCIENCE

ROCHESTER INSTITUTE OF TECHNOLOGY

ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

MASTER'S DEGREE THESIS

The Master's Degree Thesis of Gaurav Sheth
Has been examined and approved by the
Committee as satisfactory for the
Thesis required for the
Master's degree in Color Science

____

Dr. Susan Farnand, Advisor

____

Dr. Mark Fairchild

____

Date

THESIS RELEASE PERMISSION

ROCHESTER INSTITUTE OF TECHNOLOGY

PROGRAM OF COLOR SCIENCE

Title of Thesis:

Subjective Image Quality Assessment of Digitally Printed Images.

I, Gaurav Sheth, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature ____

Date

Abstract

Smartphones have become ingrained in our daily activities, driving their cameras to become better with every generation. As more and more images are being taken by cell phones it has become increasingly important to assess the quality of the images taken by different phones. While many cell phone images are only viewed electronically, many images also get transformed into printed images, especially photobooks, as digital printing gets better and cheaper compared to traditional printing processes. The gap between electronic image and printed image in shrinking rapidly and it becomes important to study the transition of images from screen to paper. The main goal of this research was to perform a rank order experiment for assessing cell phone image capture quality that translates to printed images via several different digital printers. It was of interest to investigate whether the overall image quality on displays correlates well with printed image quality. The important aspect was to study was to observe if there is a loss of image quality due to different digital printers.

Acknowledgements

I would like to thank my advisor, Dr. Susan Farnand, for her guidance, advice, and patience throughout this process. It has been a pleasure to work and learn with her over the past years.

I would like to thank my family, Hema Sheth, Aishwarya Sheth, and Amita Gandhi, who have been unending in their support. I would not be where I am without them.

I would like to thank my RIT cohort, Morteza Maali Amiri & Anku. I could not imagine navigating the program without our collaborative efforts.

I would like to thank Prof. Mark Fairchild, Prof. Roy Berns, Dr. David Wyble, and Dr. Michael Murdoch, for sharing their knowledge and time with me over these past few years.

I would like to thank Valerie Hemink for her constant help and encouragement.

Finally, I would like to thank Mojgan Rabbani and her team for providing a considerable amount of time and help in working with me on the Xerox Digital printing process.

Last but certainly not the least, I would like to express my thanks to Chanvi Kotak, whose presence helps me overcome hardships more easily and her constant encouragement keeps me motivated to achieve great things.

Contents

List of Figures

List of Tables

1. Introduction

Smartphones are becoming the most used cameras today. Camera image quality is one of the most important factors that motivates consumers towards a smart phone. The camera technology has been steadily improving every year and is at a point where it is in direct contention with DSLR cameras (albeit with some shortcomings like optics). The ease of using a smart phone camera on the go and the ability to edit and share photos at the touch of a button is what makes it a default choice for many people. Some smart phone cameras have the capability of shooting RAW images giving the user the ability to edit photos on the go with mobile applications like Adobe Photoshop, Lightroom, etc. Computational photography mitigates shortcomings that these cameras have and software gives the user control over the outcome of the final image. With enhanced image processing built-in, the end results are stunning photos with minimum input. Thus, these small cameras are capable of producing images with great dynamic range and great detail with accurate color reproduction.

Digital presses have become a power house for printing high quality images with a fast, reliable and robust print technology. They are excellent for short run jobs and can print on demand. Although the per page cost is higher compared to conventional presses, in today's digital world, demand for high volume printing is shrinking leaving digital printing as the most valid solution. This complements smart phone cameras very well as the fast image capture can be turned into high quality printed images. All this can be done inline making the entire process seamless and convenient. The capability to connect to a smartphone wirelessly

gives a massive advantage over conventional processes and automation makes it easy for a single person to handle the press, cutting on labor costs. Smartphone cameras can produce press ready images although some modifications may be required during the prepress process for ease of printing. The question that arises now is how does cell phone camera quality translate to digital press printing? Do the press parameters affect the perception of the images? The simplest way to answer these questions is performing a psychophysical study of printed images taken from various cell phones and printed on different digital printers.

2. Literature Review

2.1. Conventional Printing

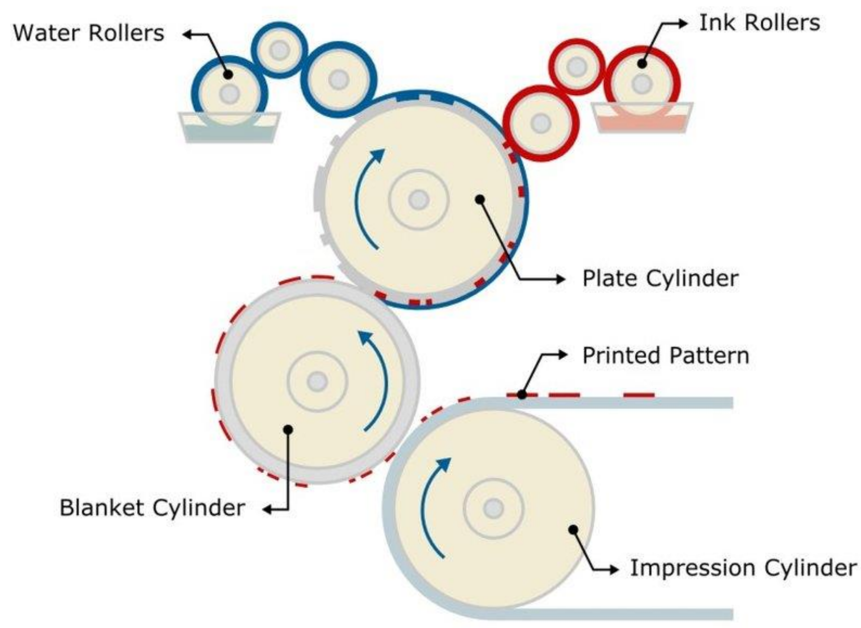
There are several methods of conventional printing that exist today. By far the most successful are Offset Lithography, Gravure & Flexography. Though these processes vary in their method of printing, the underlying principle remains the same. All processes use some form of mechanical interface to transfer ink on to a substrate.

In offset lithography a printing plate made from aluminum with the image etched on the surface that will be printed. Ink is transferred by the image part while the non-image part repels the ink. This is achieved by applying a coating of water to the plate and the ink has an oil base. The inked image is then transferred from the plate to a rubber blanket (hence the term ‘offset’) and then to the printing surface. It is mainly used in the printing industry to print newspapers, magazines, brochures, stationery and books.

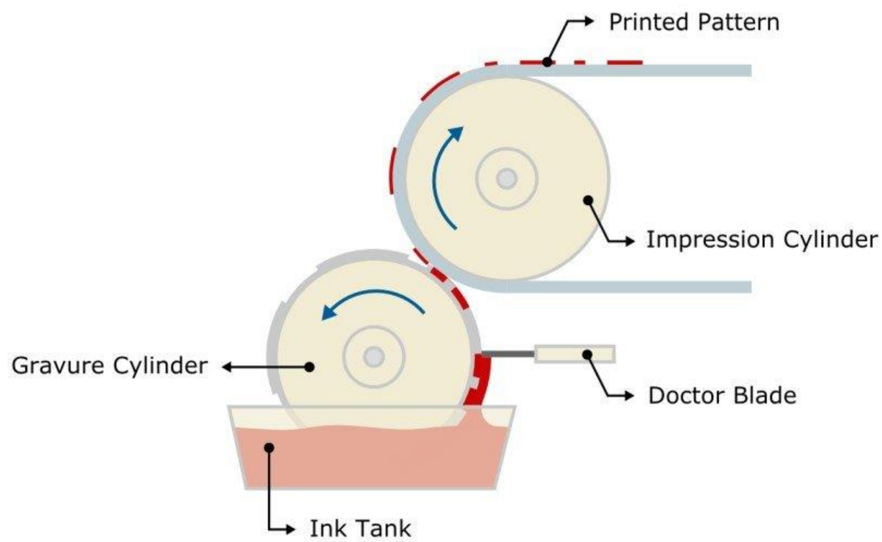
Flexographic printing process operates very similarly to the offset process with the key difference being that the image is etched into a rubber printing plate resulting in a relief printing process. In other words, the image area is raised above the non-image area. This plate is inked directly without a need of a blanket and that image is then transferred to the substrate by application of rolling pressure. It can be used with wider range of inks and can work better with water-based inks making the process easier. It’s mainly used to print on materials like corrugated

boxes, plastic, foil and rough paper and flexible material like thin films which are typically used for packaging.

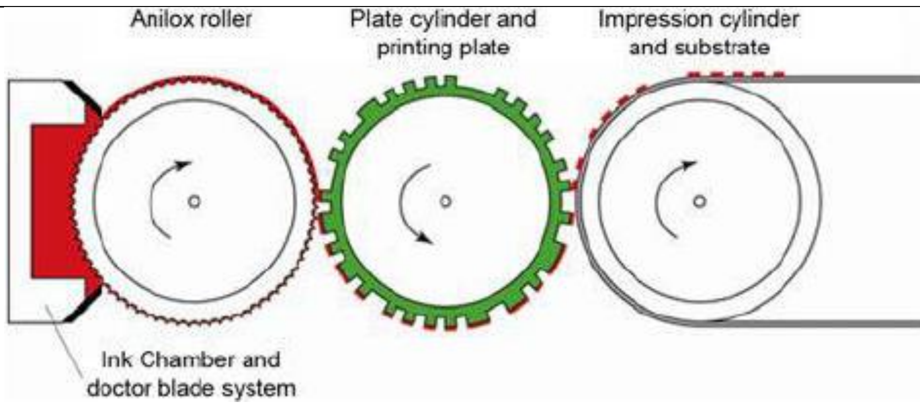
Gravure is the simplest albeit the most time consuming to set up as the image area is directly etched on the plate cylinder. The plate cylinder is dipped into the ink fountain which carries ink and transfers on to the substrate passing it through the nip pressure created by the plate cylinder and impression cylinder. It is capable of producing high quality printing as it can transfer more ink on to the substrate creating deep shadows and rich highlights. Hence it is a default choice for printing fine art and photography related material, long-run magazines (more than a million copies), wallpapers and laminates.



Lithographic Offset Printing Process [1]



Gravure Printing Process [1]



Flexographic Printing Process [2]

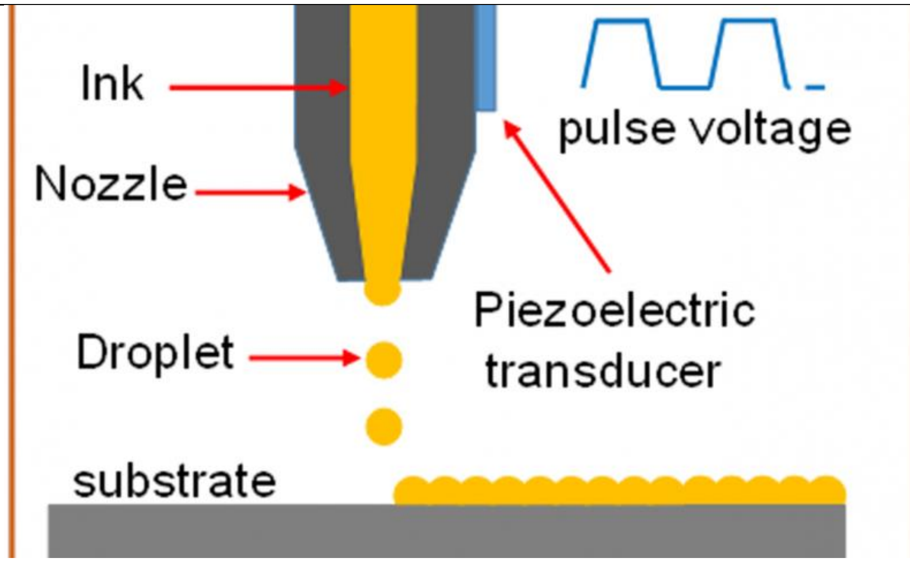
Figure 2.1-1: Conventional Printing Processes

Each process has its advantages and applications but all have one common requirement that the amount of printing needs to be large scale and in high volumes otherwise the process is not an economical method of printing.

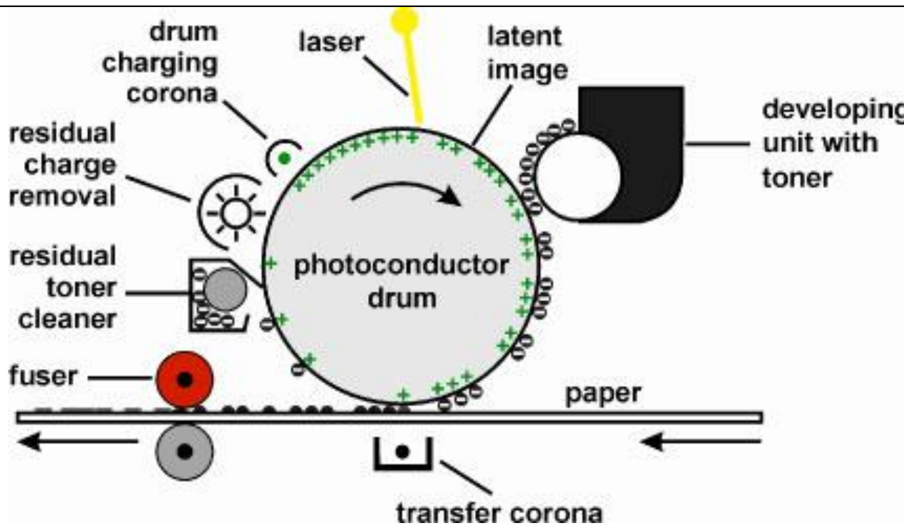
2.2. Digital Printing

Digital printing is distinctly different from conventional printing processes. The term digital printing can be applied to any printing methods that uses electronic files to produce an image using ink or toner which is directly applied on the surface of the substrate. This method eliminates the use of any setup processes like plates resulting in a fast turnaround time and on the fly modifications to the image.

There are mainly two types of digital printers, inkjet printers or toner-based printers. The inkjet printing process produces the image by applying micro-droplets of ink, ranging in sizes of 50 to 60 microns (one micron being 1/1000th of an inch) onto the paper (although newer technologies make it possible to print on film and other substrates) [3]. Toner-based printers generally use lasers to generate ionic charge on the substrate. The charge is transferred on to the substrate by a heated drum which forms a charged image on the paper. Toner consisting of minuscule plastic-based powder, which has the opposite charge, is attracted to the image area of the paper and is heated to the melting point to adhere it to the paper. The dots created by the toner-based printing process are very small ranging between 7 to 10 microns in size which provides higher resolution as compared to the inkjet printers [4].



Inkjet Printing Process [5]



Toner based – Laser Printing Process [6]

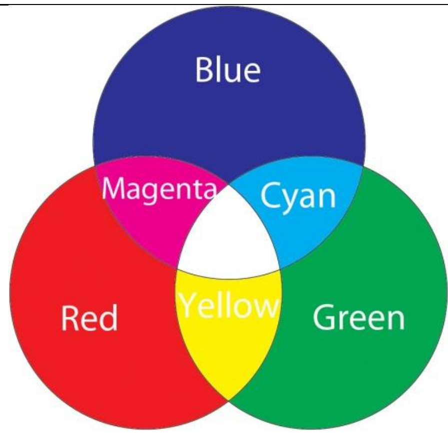
Figure 2.2-1: Digital Printing Processes

Although there are certain advantages of digital printing, its cost effectiveness lies in short run jobs which need frequent changes. Conventional process have a higher preparation cost but yield less costly printing for long run jobs with few to no changes. As today's world is mostly digital the need for on

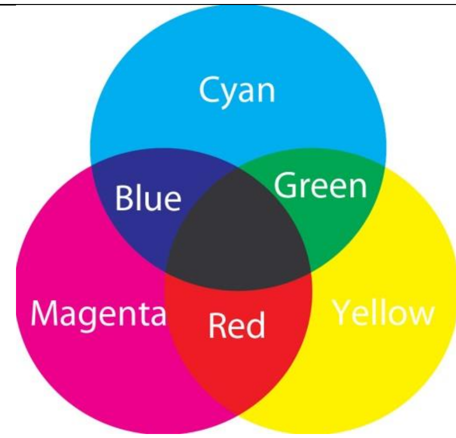
demand printing has risen significantly. With advent of new technologies and research in digital printing the cost gap is shrinking rapidly to the point that it has essentially disappeared and it would be safe to say that because of the versatility and high quality, digital printing will outshine conventional printing processes.

2.3. Color reproduction.

There are two color theories that are predominant in reproducing color. One is the additive color theory based on primary colors of light – red, green and blue. As shown in Figure 2.3-1, the additive reproduction process mixes various amounts of red, green and blue light to produce other colors. Combining additive primary colors with other primary colors produces the additive secondary colors cyan, magenta and yellow. Three primary colors produce white when mixed in equal proportion. Monitors, digital cameras other such devices work on principle of additive color theory [7]. The other theory that is widely used by printers is the subtractive color theory. Additive secondary colors constitute the primaries for subtractive color theory, which is demonstrated in figure 2. When mixed yellow and magenta give red, magenta and cyan give blue, yellow and cyan give green. Printing is done by using these secondary colors and an overlap of these colors produces a full gamut of colored image [8].



Additive Color Theory



Subtractive Color Theory

Figure 2.3-1: Color mixing theory [9]

Printing is usually done on a white substrate by using a set of process colors YMCK, the subtractive primary colors, plus black. Black is added to cyan, magenta & yellow as it gives more depth to the image. Combining cyan, magenta and yellow will produce black theoretically but in reality, they produce more of a muddy brown as compared to black plus using three inks to produce black makes the process expensive and hence black needs to be added as a separate ink. To print on to a substrate the full color tone is broken down into halftones by separation of the colors into C – cyan, M – magenta, Y – yellow and K - black. These YMCK tints are called process colors, and they are printed on paper to reproduce the continuous tone full color image. The separation process breaks down the image into tiny dots that are printed in specific angles on the paper to reproduce the original image. All four colors have different printing angles called as screen angles that can be adjusted. The separations are used to produce four different image carriers or a more common term ‘plates’. These plates are used to print the image,

but the image is reproduced with halftone dots and not as a continuous tone image. The dots are very small ranging from 2 to 5 microns in size and to the naked eye give an appearance of a continuous tone image [10].

2.4. Raster Image Processing

The raster image processing (RIP) is the technique used to process the image areas and convert it into the dot format necessary for printing. For conventional processes the RIP computes all the areas that must be dotted so that the image will be engraved on the printing surface. Ink jet printers use similar RIP process to generate one-bit data that will enable the inkjet nozzles drive the desired color of ink on to the substrate. Thus, the resolution of 600 to 1200 DPI can be achieved to get a high-quality printed image. Similarly, laser printers exploit a similar RIP process to modify the electrostatic charges. Every color in the printer has a distinct channel that charges the drum a specific spot on the drum to attract the toner to that proper location. Thus, the rasters work in unison to produce an accurately colored image [11].

Halftone screening is an important part of Raster Image Processing. Halftone screening enables the image data being printed to have grayscale and tonal range. RIP converts the image data to a matrix of bit data, but that’s not enough to print a good colored image. Halftoning translates the eight-bit tonal values into one-bit data by dividing the area into dots of different shapes and sizes producing a smooth gradient. Therefore, RIP is a two-step process, converting RGB data to CMYK continuous tone data and converting continuous tone data to

halftone dots to produce a vignette [12]. Vignette is defined as a continuous tone image produced by printing dot pattern.

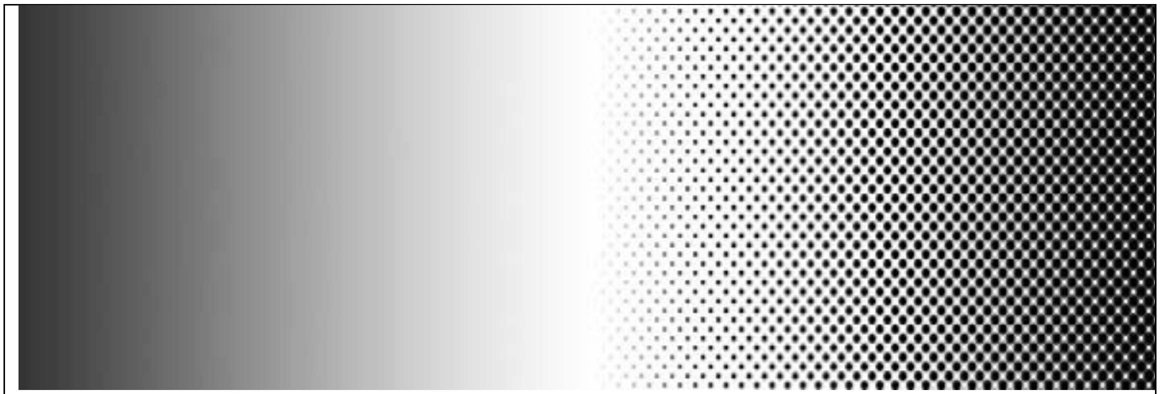


Figure 2.4-1 – Vignette pattern created by printing dots of varying size and spacing

Thus the understanding of the ripping process is important to understand the limitations of the printer and the expected image quality from the said printer. Raster Image Processor can only convert given data into raster for printing. It cannot interpolate missing data to generate details in the image so it is of paramount importance to have a high-quality image with good detail as the starting point of the printing process.

2.5. Significance of Digital Printers

Digital printers have the ability of printing from a digitally created image straight on to the desired media without the hassle of extensive setup. Digital printers are ideal for running low volume jobs. Large-format electrophotography/inkjet printers are typically used to print as they are standard

methods and have been established over the years. Such printers usually have higher cost per page as compared to conventional methods but they make up the cost by avoiding the setup cost of plates and the ability to print on demand thus negating storage inventory cost. Additionally, they can print faster saving precious print time and have the ability to customize the image when required during a running job [13]. The image quality and printing capabilities of digital presses are improving rapidly to the point where they can compete on every level to the conventional printing capability to print large scale jobs at a low price [14].

Most significant advantages of digital printing are:

- Digital files used on the press are simpler to edit and modify as compared to the files required for plates. Changing plates is cumbersome and time-consuming
- Initial setup time and cost is minimal
- The presses have laser guided electronic color registration system built in as default.
- Digital presses are efficient in proofing and are excellent for short run jobs
- The input images can be changed as required giving digital printers the capability to do variable data printing, personalize the printed job, print on-demand, and print directly from the web.
- Provide flexibility in printing by collating jobs that require multiple operations by using automated in-line processes.

- Automation of the process eliminates the need for multiple operators as a single operator can handle the job from start to finish. This reduces labor cost and increases efficiency.

2.6. Image Quality

Print image quality refers to the general appearance of the print including the presence or absence of defects detected by the observer that may result in rejection of the printed sample. Assessment of print quality attributes determine if the printed job satisfies the required quality criteria. ISO standards allow print quality metrics to be quantified and used for quality control purposes [15].

The following print attributes are crucial to assess in evaluating digital print quality:

- Sharpness – Sharpness is defined as number of dots printed per inch. For example, printing 2880 dpi results in sharpness equivalent to 1440p.
- Gloss – It is determined the substrate used rather than printing process or inks used. Glossy paper has a smoother surface providing better print quality.
- Color / Tone – Color accuracy and tone mapping are required to produce good color image with accurate skin tones produced by good tone mapping.
- Banding – Smooth print lines are broken into jagged steps creating variations in the printed image.
- Color variation and non-uniformity – Color variation within the printed image or successive printed images. Neutral images have a tendency to have blue/yellow cast if media is not calibrated creating non uniformity.

- Ghosting – Ink residue left on to back side of successive sheets due to high ink film thickness (IFT). Also, when printing is non uniform some areas have higher IFT as compared to other areas.
- Graininess - Grainy appearance in the image area due to poor RIP, rough substrates, or charging issues, equivalent to noise in electronically displayed images.
- Hickies - Imperfections caused by dirt and debris stuck under the substrate or by coagulation of ink particles.
- Mis-registration - Poor alignment of the rasters causing the print areas to be printed in wrong spots.
- Moiré pattern – Undesirable pattern observed due to misregistration of the screen angles.
- Mottle – Non uniformity observed due to uneven ink coverage in uniform print area viz. sky, wall etc. Uneven drying occurs due to mottle causing ghosting and other defects.
- Line and text quality – low resolution causes fine lines and texts to reproduce incorrectly leading to illegible texts.

2.7. Psychophysics

Psychophysics is the study of the response of psychological phenomenon with the application of physical stimulus. It is not only a study of stimulus responses, but the study of the methodology used to produce the stimulus as well. It is a

powerful tool that can be used to measure the perception of observers and record their performance [16].

The response of subjects to the induced stimuli has a large variance due to the influence of many factors, hence the physical stimulus used in the experiment needs to be carefully induced. If the responses are not measured correctly the results of the test can be hard to interpret [17]. Each experiment needs to be carefully planned as the parameters of other successful experiments might not work for other experiments as in there is no ‘one size fits all’ approach.

There are seven broad categories of psychophysical experiments:

- Absolute thresholds: Lowest level of stimulus that can be detected.
- Differential thresholds: Smallest measurable difference between two stimuli that can be quantifiably measured.
- Equality: Measurement of two physical stimuli with differing attributes but can be perceived as equal.
- Rank ordering: Arranging a set number of stimuli in an ascending or descending manner according to their perceptual attributes.
- Equal interval: Physical stimuli that produce equal differences in some perceptual attribute.
- Equal ratio: Physical stimuli produce equal ratios in some perceptual attribute
- The subjective estimation of physical attributes [18].

These seven categories have vast number of applications. The reliability and accuracy of the categories varies a lot, so it is crucial to select he proper method for the experiment. For the current study rank order is chosen as it is fast, and the number of input stimuli are large. Hence, we will focus on rank order in the next section.

Rank ordering is one of the simplest methods of ranking samples in an ascending or descending preference of quality. The rank order experiment usually tasks the observer to rank the stimuli, usually focusing on one attribute at a time. Thus, if there are n number of samples the rank order will be from one to n. The idea is to get rankings from M observers, then assign a number to the ranks and calculate the average of the ranks. The important point to note is that this method only gives information about the sequence of the samples that have greater amount of a particular attribute as compared to the previous sample [19]. There are two ways for data collection for a rank order experiment, both produce identical scale values but represent the data in dissimilar fashion. First method employs the formation of matrix R with a column indicating each sample from 1 to n. Observers are represented in the first column. For every observer, the rank specified for the observer is recorded for the stimuli indicated by the column. Thus, the data matrix R is J rows X n columns. This method is suitable for converting the data matrix into proportion matrix and is usually the preferred method. From the proportion matrix, an interval scale can be easily obtained [20].

	Sample number				
Observer	1	2	3	…	n
1	3	1	N	…	6
2	2	3	n-1	…	5
3	3	1	n-2	…	6
…	…	…	…	…	…
J	4	1	n	…	7

Table 2.7-1: Example of data matric R form a rank order study

$$AvgRank_1 = 1/J[1\ 1\ 1\ \ldots\ 1_j]R \ldots \text{(Eqn 1)}$$

The second method gathers the data as a histogram matrix. In this method the rows are ranks 1 to n. Each block of the data matrix H consists the number of times the sample was judged to have the rank specified by the row. Data matrix H should have as column sums the number of times each sample was ranked. Thus, the data matrix H will be in a representation of a Histogram, with a row for each rank, a column for each sample and cells containing the frequency of rank selection. Dividing each entry by the number of times the sample was ranked will yield an approximation of the probability of the sample being assigned the rank. The average rank can then be calculated by multiplying the probability by the rank and calculating the sum [20].

	Sample number				
Sample Rank	1	2	3	…	n
1	10	0	0	…	6
2	31	2	0	…	1
3	27	27	1	…	5
…	…	…	…	…	…
r	0	1	5		0

Table 2.7-2: Example of Histogram data matric H form a rank order study

$$AvgRank_2 = 1/J[\ 1\ 2\ 3\ \dots\ n]H \dots \text{(Eqn 2)}$$

It should be noted that both methods only give the order of the ranking and not the interval scale. The difference or distance between the samples that are ranked cannot be calculated. Thus, the numbers generated as a result of calculation are not significant and do not convey interval between the samples.

The goal of this study is to perform a rank order experiment for assessing print image quality of several different digital printers. It was of interest to investigate the overall image quality that each of the presses produced. The important aspect was to study if there is a loss of image quality due to different digital printers. Rank order is a good psychophysical tool for assessing image quality perceived over a wide range of differences. Rank ordering is chosen over paired comparison, as the number of pairs required to be compared would be large and the experiment would become long and cumbersome. As the image quality is

high and it is not easy to distinguish between the images, there is bound to be confusion among the observers. If there was minimum confusion, a direct scaling method would have been a better approach; since confusion was predominantly present, an indirect approach was chosen. Once the observers were shown the images their response was used to estimate the scale values along the desired dimension. This information was useful for figuring out the scale value which signify the reliability of the choices made by the observers.

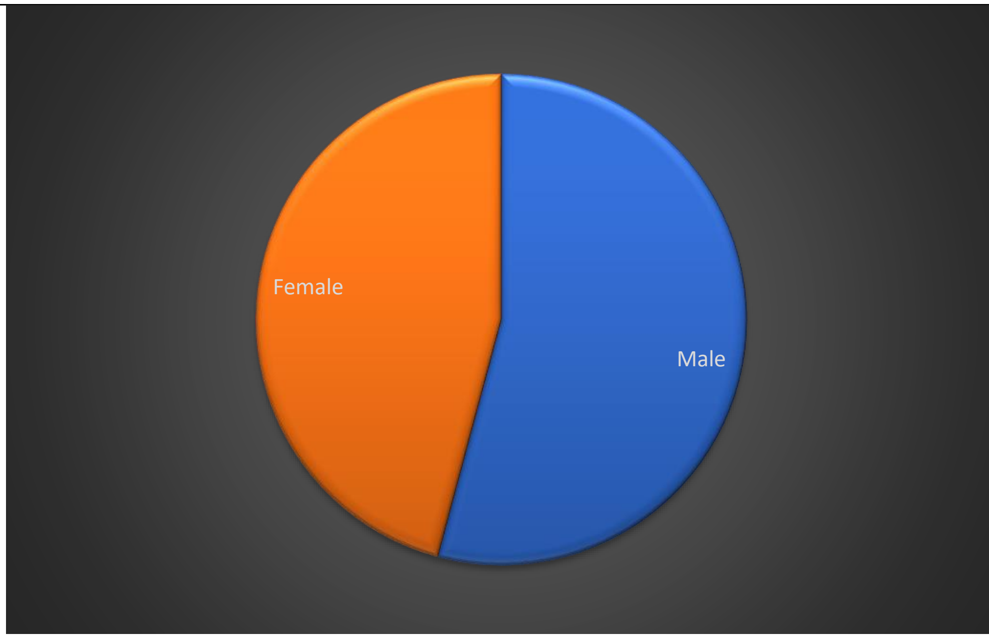
3. Experiment 1

Experiment was designed to assess the difference in perceived image quality of different digital printers. Images from different digital printers were rank ordered by observers according to their preference. It was crucial to determine the preference of digital presses to gauge their impact on the print quality and their viability in the next experiment. The key element was to study how the printers the produce the best image quality given the optimal settings for each printer.

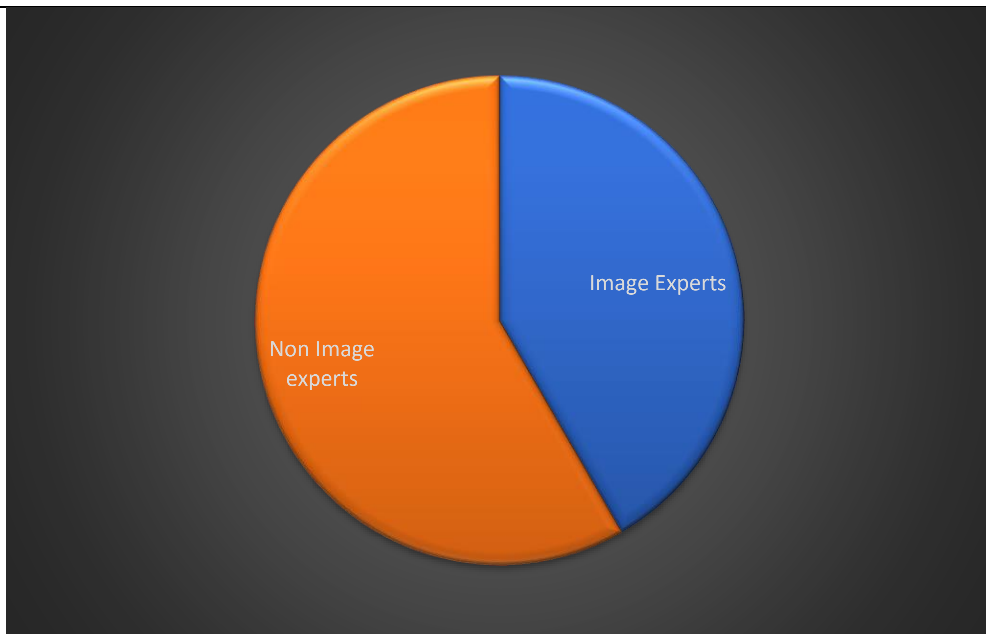
3.1. Method

3.1.1. Subjects

Twenty-Four observers participated in the Image Quality Experiment. Figure X describes the gender, age and field distribution of observers. Thirteen of observers were male and 11 were female. Different age groups of subjects were incorporated however a major portion of the observers were students from RIT. Apart from different age groups people with imaging background were included to gauge the difference of their observations against people without an imaging background. This is important because people working in the imaging industry have an eye for noticing minor defects in the prints whereas for others the image might look perfectly fine.



a



b

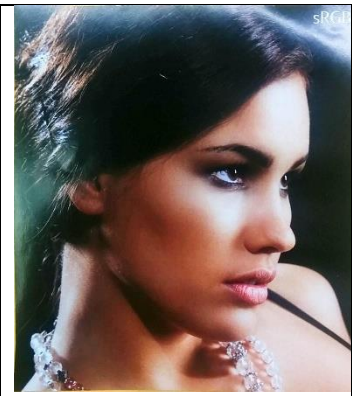
Figure 3.1.1-1: Observer Distribution: a. Gender & b. Background

3.1.2. Apparatus

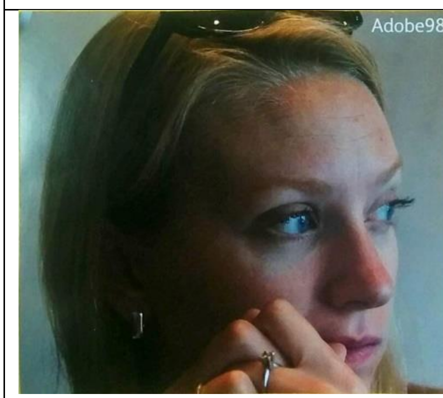
This study included generating the image quality ratings associated with real-world test images using rank order protocols. Rank order is a good psychophysical tool for assessing image quality differences perceived over a wide range of differences. The images were printed by 7 different commercially available digital printers with a resolution of 2400 DPI running 100 pages per minute with a maximum media size of 13" x 19.2" (330 mm x 488 mm). The images had embedded ICC color profiles that were honored by each printer. The media used was selected based on the capability of the machine to produce highest quality prints. To keep the comparison valid, the press operators were strictly instructed to use media between 200 – 250 gsm. The slight variation in the grammage was unavoidable as some presses need a specific gauge paper to produce the best overall image quality. A total of 8 images were chosen for the experiment. A subset of the entire image was cropped so that the observer will not have other clues during the process (Fig 3.1.2-1). Each subset was chosen to have a different attributes such as gloss, color uniformity, skin tones, graininess, sharpness etc. that varied among the printers. The experiment was conducted in the Munsell Color Science Lab at the Rochester Institute of Technology.

1 – Gray Balance

2 – Highlight Detail



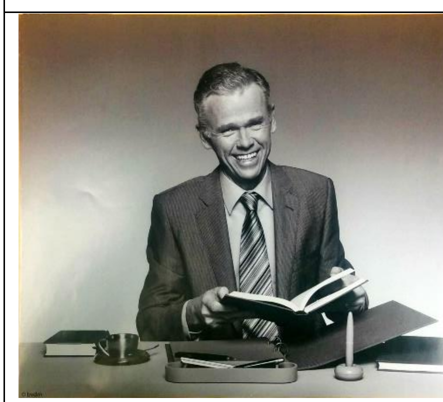
3 - Flesh Tones (sRGB)



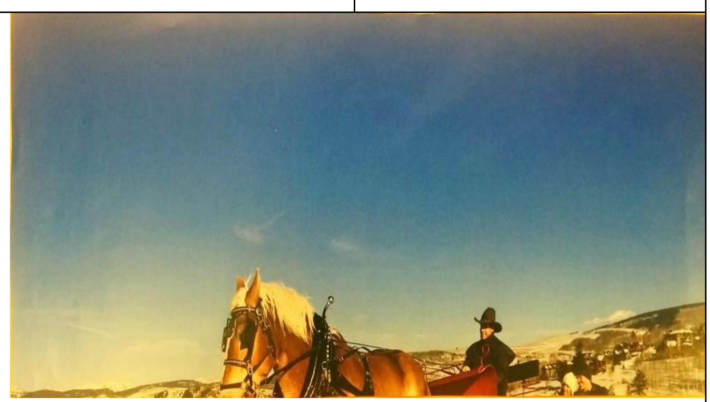
4 – Flesh Tones (Adobe 98)

5 – Flesh Tones (sRGB)

6 – Shadow Detail



7 – B/W RGB/8



8 - Graininess

Figure 3.1.2-1 : Images used for the Experiment.

3.1.3. Procedure

Observers were instructed to relax and sit in front of a X-rite light booth. Samples were randomized before handing over to the observer. The observers were asked to wear gloves in order to not damage the printed samples. The light source chosen for the experiment was D50 standard light source. The observers were asked to rank order the images according to their preference and stack them in order of most preferred to least preferred.

Figure 3.1.3-1 – Viewing setup for the experiment.

Following are the instructions given to observers:

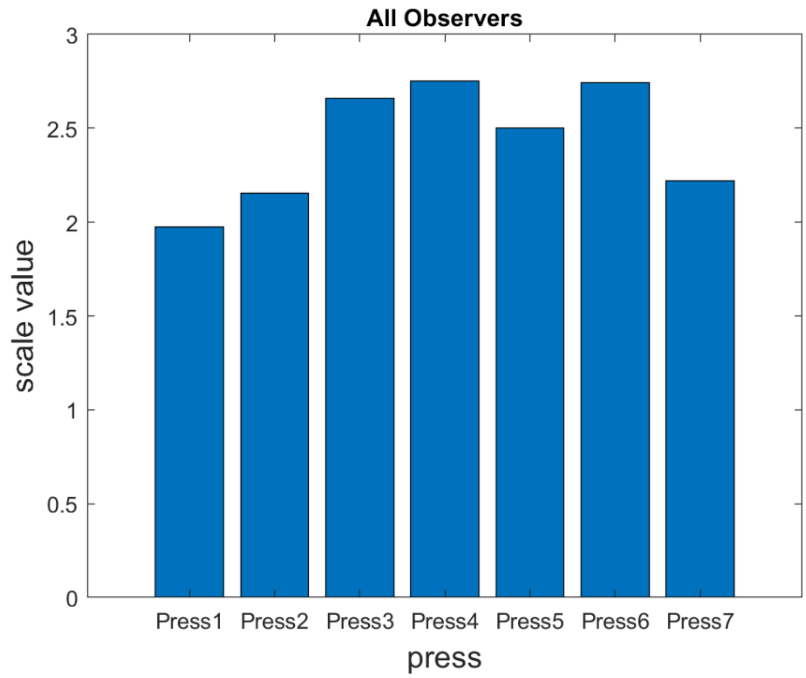
"Print image quality is divided into several different categories. Each category affects the overall perceived image quality of the printed sample. In this experiment, you will measure the effect of each parameter on the image quality. You will be given a set of similar images that are printed using different digital printers. Your task is to rank order each set of images according to your preference. You may lay all images side by side in the light booth and then select your order. Order the images from the most preferred to least preferred. You can stack them on top of each other."

3.2. Results and Discussion

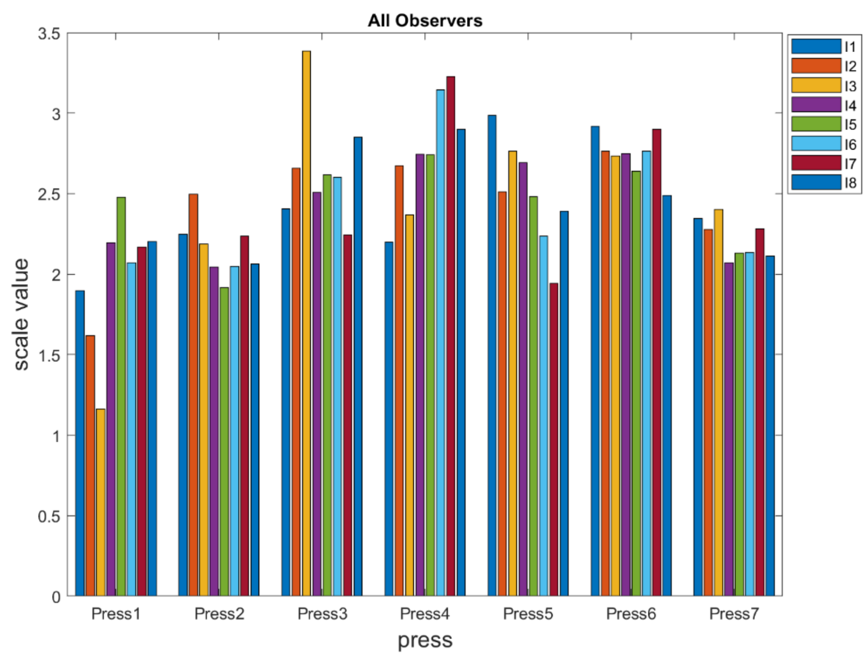
The rank order data was gathered in an Excel® file. The back side of each print had the press name and embedded profile stamped but the observers were unaware of this fact so as to have them rank the prints according to their preference and avoid any bias. The data contains 7 columns; each column has a preference rating that the observer provided for a particular image set. The data set has 8 rows with each row containing the specific parameter ranked by the observers. A variety Flesh tones (Skin tones) were chosen as people tend to prefer different tones according to their ethnicity. The data is analyzed using Thurstonian method V. The rating for each observer is converted to a frequency matrix, which was then converted to a proportion matrix. The proportion matrix was summed and converted to z-score which is adjusted to be all positive by adding a value of 2.5 for ease of viewing.

		Press1	Press2	Press3	Press4	Press5	Press6	Press7
Gray Balance		7	4	5	6	1	3	2
Highlight Detail		6	4	7	3	2	1	5
Flesh Tones	Image 1	7	6	5	3	2	1	4
	Image 2	5	4	7	3	1	2	6
	Image 3	5	6	3	4	2	1	7
Shadow Detail		7	5	3	6	2	1	5
B/W RGB/8		7	5	3	6	2	1	4
Grainines s		2	3	5	7	6	4	1

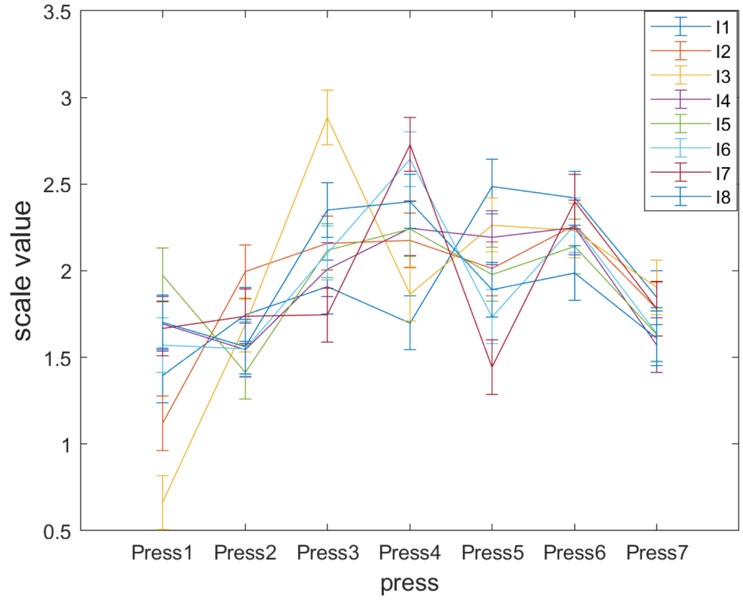
Table 3.2-1: Response of one of the observers for Press evaluation



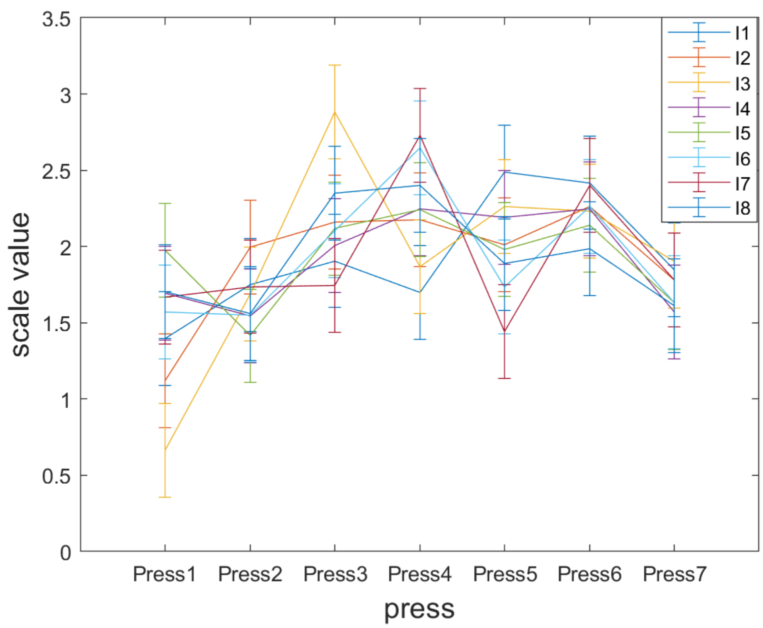
A – Average



B – Individual Images

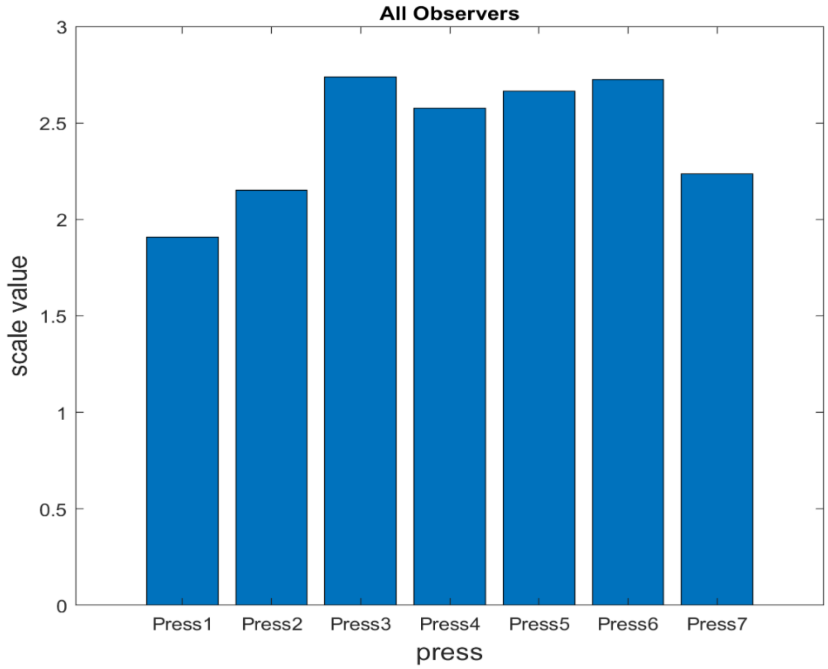


C – Error Bars

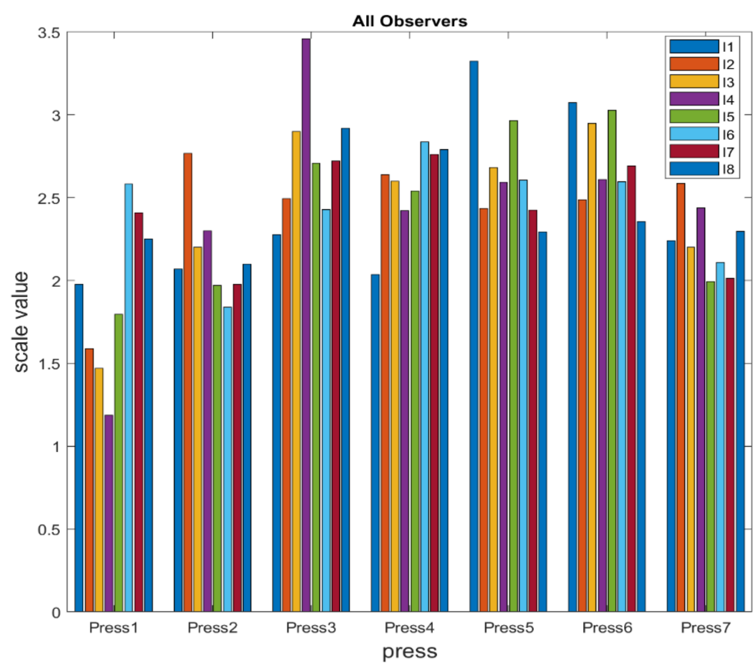


D – 95% Confidence Intervals

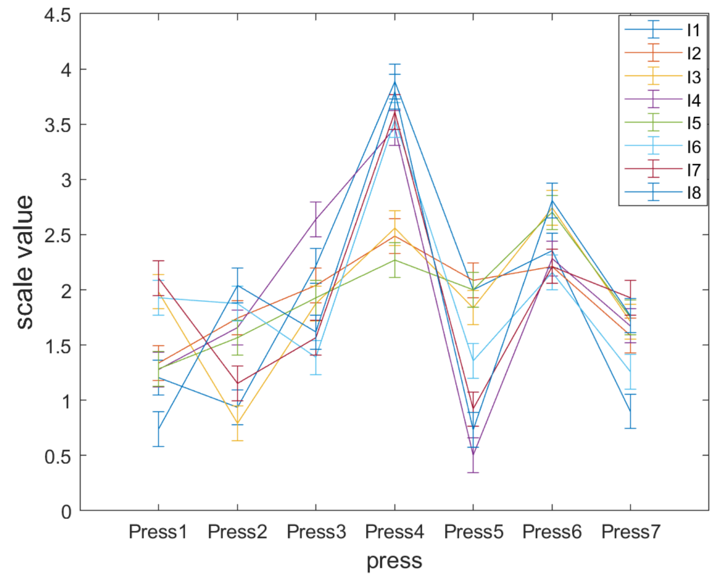
Figure 3.2-1: Preference plots for all observers



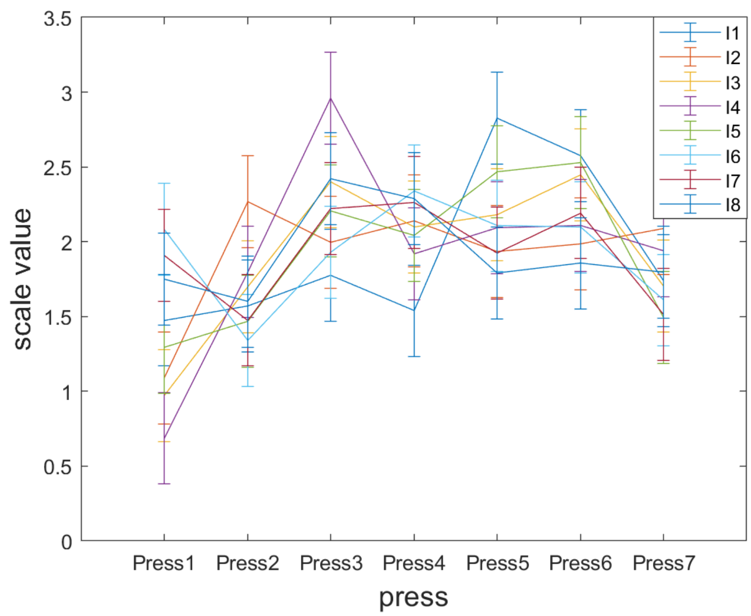
A - Average



B – Individual Images

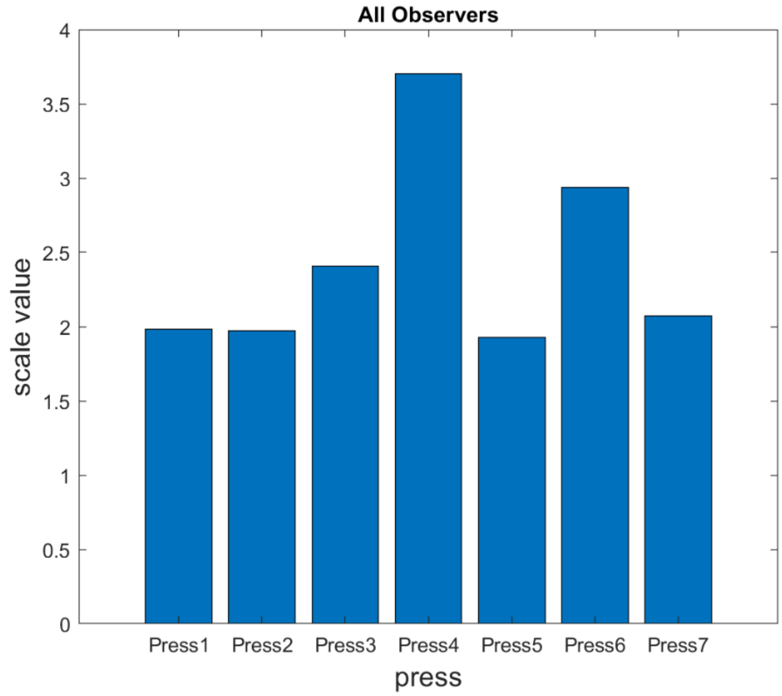


C – Error Bars

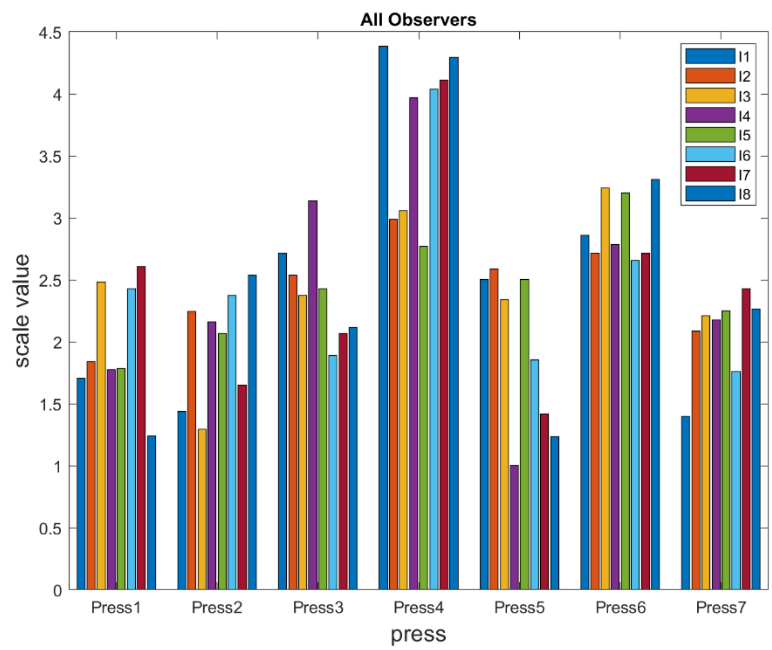


D – 95% Confidence Intervals

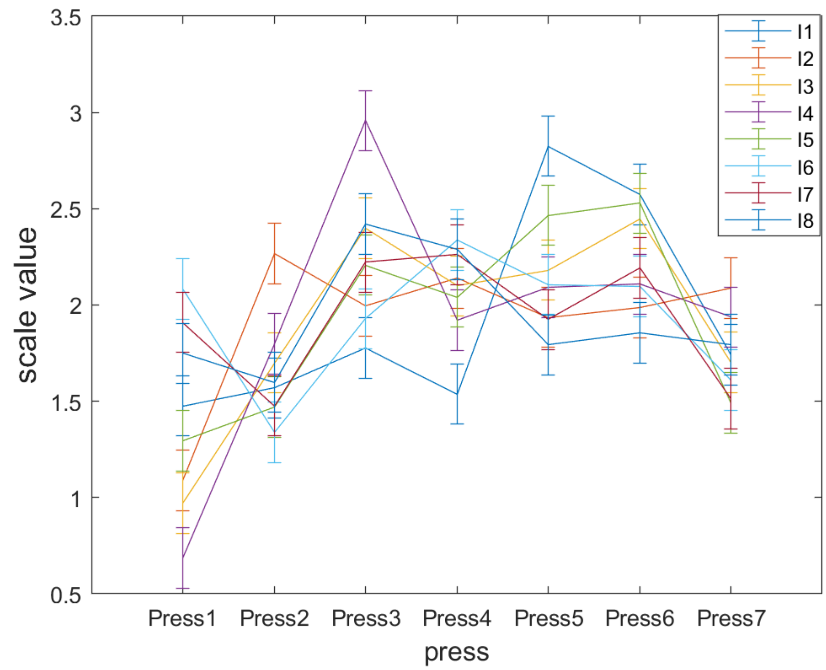
Figure 3.2-2: Preference plots for non-image experts



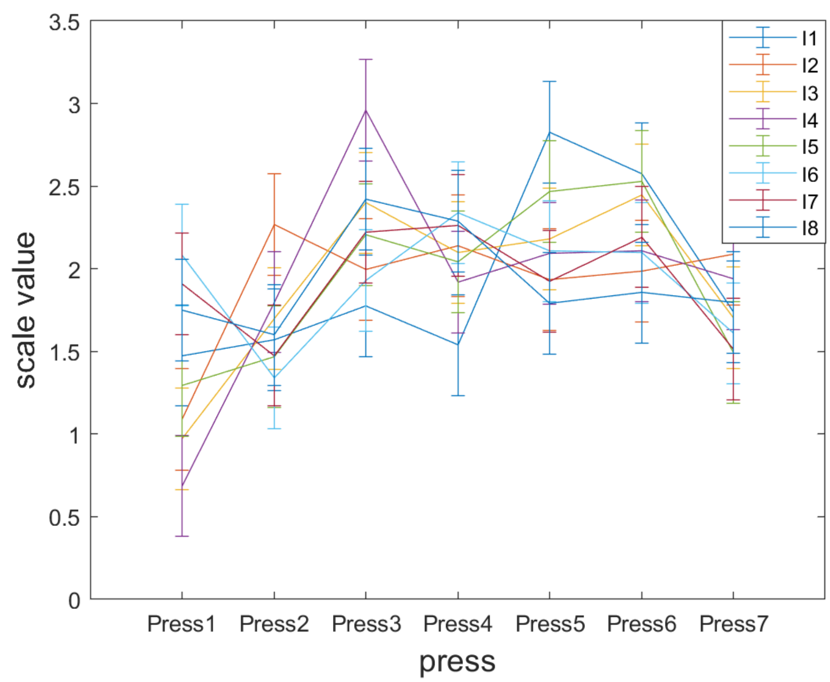
A - Average



B – Individual



C – Error Bars



D – 95% Confidence Interval

Figure 3.2-3 : Preference plots for Image experts

Every digital press has capabilities that produce minor differences in the print quality. Small differences in the RIP create major improvements in the image quality. Also calibrating the press periodically certainly helps in improving print quality. Figure 3.2-1 shows the overall preference of all observers. The image quality of all presses is very close, and it is hard to distinguish between the prints. Presses 3, 4 and 6 have a slight edge over other presses but as is seen in Figure 3.2-1 but the difference is minimal with a scale value difference of less than 0.5. Press 3 & 4 and Press 6 are direct competitors and the observers had a really hard time picking between them. The error bars show the confusion of the observers very well and it can be seen that the observers could have leaned either way when choosing the order. The Confidence interval shows the likelihood of the observers choosing the same images in the same order and hence it is observed that CI are larger than standard deviation. The likelihood of the observers choosing the exact same order is rather low but the end result obtained by averaging the overall preference of the observers viz press 3, 4 & 6 yields a consistent result. As more and more observers participate in the study, it is possible that the difference in the presses would get larger. The additional observers would provide more robust results.

People working in the print and imaging industry (termed as image experts here) have an eye trained for looking for defects. People in general (termed as non-image experts here) tend to neglect or rather not notice small defects as for them the impact of those defects is not large. Comparing the plots from figure 3.2-2 & figure 3.2.3 shows a stark difference between the results. Image experts found

small defects in the printed samples and thus have a high preference for press 4. The error bars as well as the confidence Intervals are smaller. The non-image experts seem to have a lot of confusion in choosing one print over other and thus all the preference seem to level out. This also increases the error bars and Confidence intervals as there is a high chance they will change their preference in successive runs. The strong preference of image experts also skews the preference plot of all observers combined thus providing more real-world results. Hence it is important to have a mix of image experts and non-image experts when performing such an experiment. The results for image experts are important because, once observers are trained, as customers might be, they can find these defects easily.

4. Experiment 2

This study employs an indirect scaling method for rank ordering cell phone images. As the image quality is high and it is not easy to distinguish between the images there is bound to be confusion among the observers. If there was minimum confusion a direct scaling method would have been a better approach but since confusion is predominantly present an indirect approach was chosen. Once the observers are shown the images their response is used to estimate the scale values along the desired dimension. Ranked order data gives a lot of information on the observers' preferences for the set of input stimuli. The information is useful for indicating scale of dimension and the reliability of the choices made by the observers. Thurstone presented a theory of the discriminal process (the process by which we make judgments of samples) and how its nature allows us to construct an interval scale based on comparison of stimuli. The methods developed by Thurstone to analyze simple rank ordered data makes use of the confusion of the observer data and offer improved tactics for data analysis over simple mean preference.

4.1. Methods

4.1.1. Subjects

Twenty-four observers participated in the experiments in the Visual Perception lab in the Color Science Hall at RIT. The observers ranged in age from 23 to 64, included 10 males and 14 females, and included 10 Color Scientists and 6 people

from imaging backgrounds and the rest did not have imaging experience. For the printer with image enhancement only 11 observers took part in the experiment.

4.1.2. Apparatus

The study involved use of eight phones viz. Galaxy S7 Edge [21], HTC One M8 [22], iPhone 4, iPhone 5S, iPhone 6S Plus [23], an LG G2 [24], Nokia 1020 [25], and Sony Z5 [26]. Four digital presses were used to print the images used in the study with modifications to the print settings on one of the presses. The presses involved were IGEN, IGEN with color enhancement, Color press 1000, Versant 2100 [27] and Shutter fly [28]. For the sake of convenience IGEN with image enhancement is considered as a separate press in the study. The color enhancement is obtained by changing the image processing algorithm to improve contrast and boost color saturation. The device cameras were evaluated in a lab setting before they arrived at RIT and were used in the experiment. The main goal was to perform rank order study of the printed samples. This was achieved by capturing a variety of scenes with each of the phones and printed on different digital presses. The printed images were shown on similarly specked sheet of paper so as to keep the method of viewing consistent.

To display similar fields of view, phones with larger pixel heights needed to be moved farther from the target. A look-up table (LUT) was generated for each to give an idea of the distances required between the target and the phone. The pixel height of an object in the field of view was measured in each image; the images where the pixel height of the object was consistent between phones were then

chosen. The settings under which the captures were made were carefully controlled. The phones were mounted on a tripod to keep them stable and the flash was turned off. The level of illumination at each target was measured with a Minolta CL200 lux meter and recorded.

	Scene	Illuminance (lux)
1	Empty Restaurant	257.8
2	Flowers – Blue LED	731.7
3	Flowers – Outdoors	>99,999
4	George Eastman House	14450
5	Handicapped Sign	>99,999
6	House – Night	15.9
7	Person in Garden	>99,999
8	Portrait - Fluorescent	1150
9	Portrait – Low Light	15.4
10	Portrait - Outdoors	>99,999
11	Portrait - LED	140
12	Sign – Night	16.1

Table 4.1.2-1: Illuminance level for each image scene.

The settings were carefully controlled, and the phones were stabilized by mounting them on a tripod. The image scenes were chosen in in order to cover a wide range of targets and illumination. The scenes included buildings, flowers indoors and outdoors, fruit scene, George Eastman building, handicapped parking sign, House at Night, people under low light, fluorescent lighting and outdoors.

Figure shows the images that were chosen for the study. There was some variability in facial expressions between images of the models taken with the different phones. It has been proven by other studies that people are more likely to rate an image of a person smiling more positively than an image of a person frowning. Because of this, participants in the psychophysical experiments were given specific instructions to ignore the facial expressions of the models and to judge the image only on overall quality but it is unknown how well they were able to comply with these instructions.

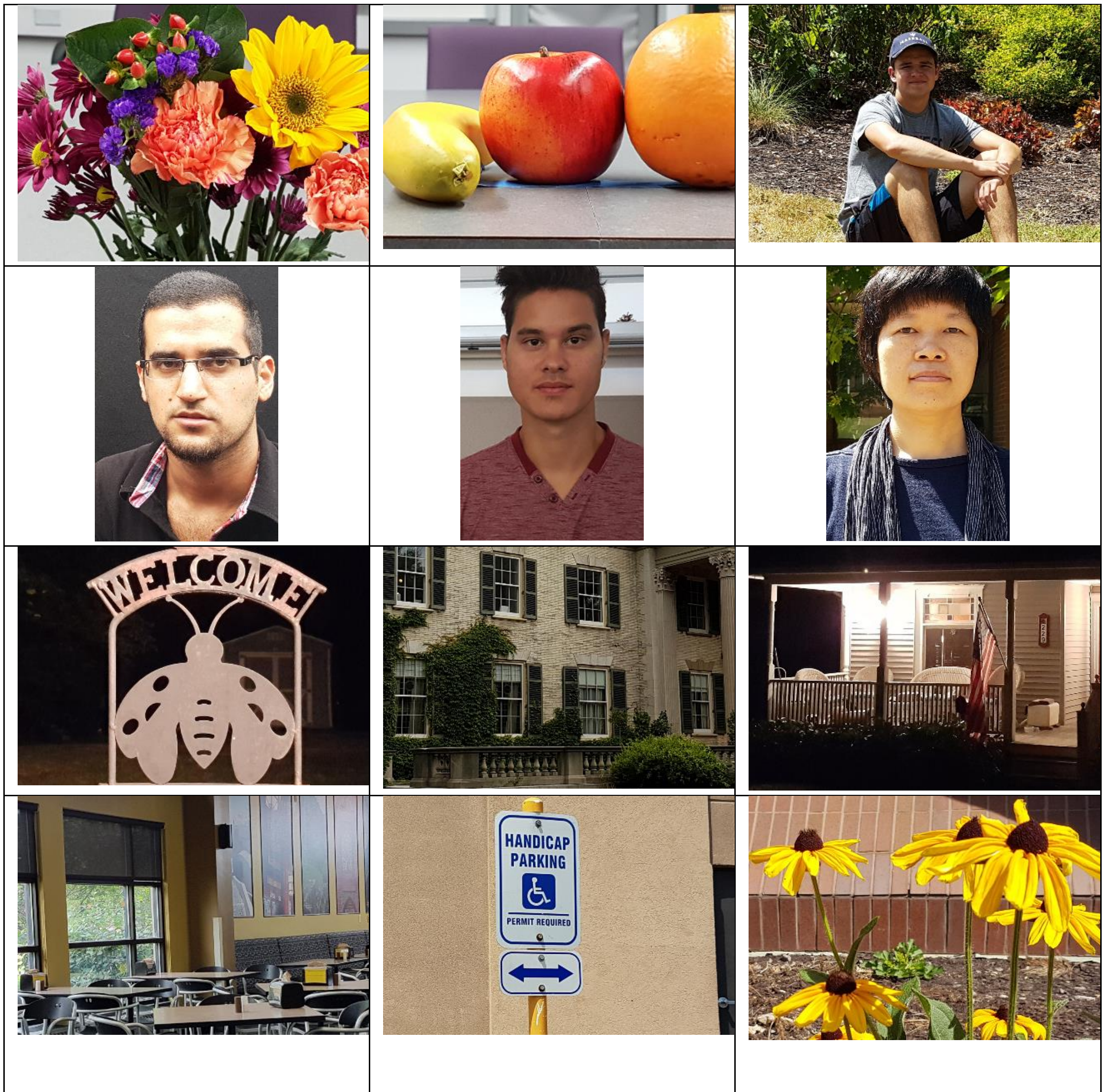
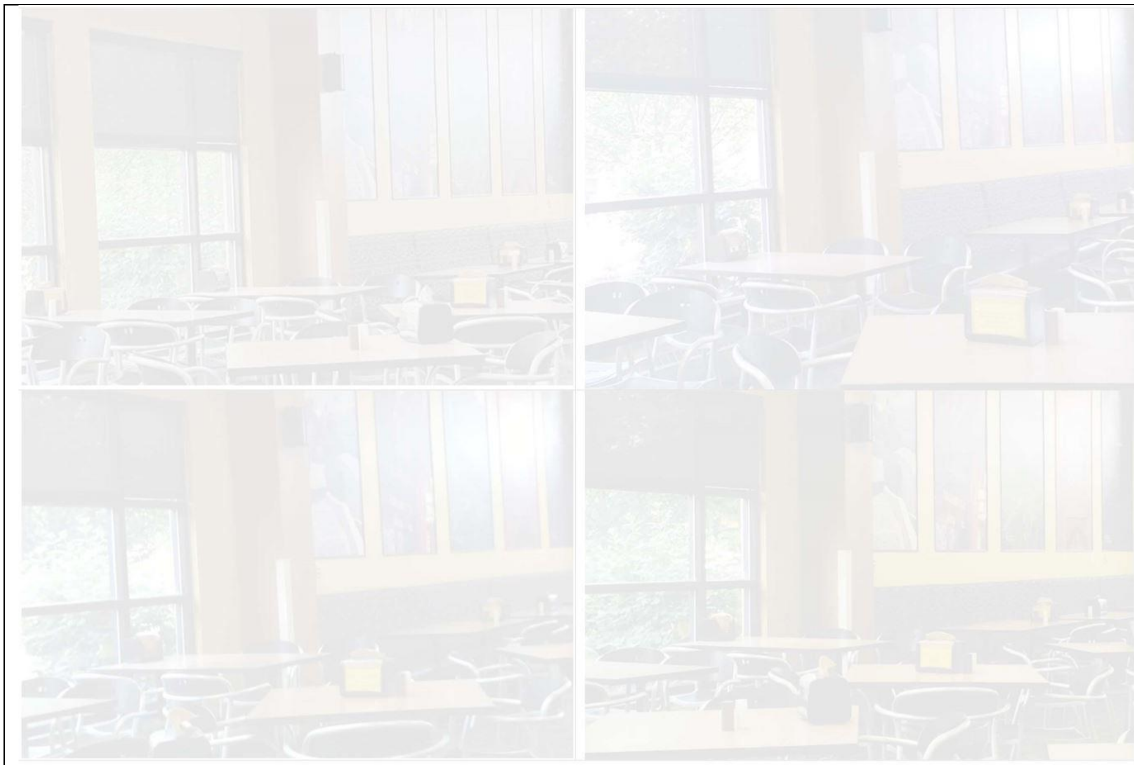


Figure 4.1.2-1: Images of different scenes chosen for the experiment

The presses that would produce the best image quality were selected. Shutterfly® is a commercial press, the Shutterfly® layout was used to resize the images to be of the same aspect ratio that is available on their website [28]. For the other presses (IGEN, Color press 1000 & versant 2100) Adobe InDesign® was used to make the raw files of the different images with same aspect ratio as Shutterfly®. The page size used was 8.5 X 11 in landscape mode. Each page

contained four images with the size of 4.15 X 5.4 each. The paper used for all presses was 200 gsm coated stock. Images were put in randomized order and each page had 4 images of the same scene. So, two pages contained all 8 images of one scene. The background was chosen to be neutral gray across all prints.

Example of the arrangement of images on each page



Example of transparency page of the answerer sheet.

Figure 4.1.2-2: Experimental setup for experiment two

4.1.3. Procedure

An answer sheet was provided to the observers that were exact copies of the images being viewed. The images on the answer sheet were 1 percent transparencies of the images that were created in Adobe Photoshop. The observers were told to write their preference for each image on the answer sheet going from 1 to 8 with 1 being the best preferred and 8 being the least preferred.

For the experiment, the first page presented instructions for the participants as follows:

- *Please rank order the images by your preference in the provided sheet.*
- *The scale is from 1 to 8 (1 being the most preferred and 8 being the least preferred.)*
- *Please use the entire scale.*
- *No images can have the same preference*

4.2. Results and Discussion

The results contain 8 columns; each column has a preference rating that the observer provided for a particular image. The data is analyzed using Thurstonian Method described below. The table describes the Thurston analysis for the first image set. Step one shows a 8*8 matrix (for each cell phone) which is a frequency matrix that contains in each location of matrix the number of times the image of *j'th* column is chosen over the image of *i'th* column, where *i, j* are the 8 images taken from all 8 cellphones. Step 2 calculates the proportion matrix by dividing by the number of observers (N). Step 3 is converting these proportion into normal deviates (z-score) and the average of these columns defines the interval scale values for the stimuli. All the above steps are followed for all 12 images.

Table 4.2-1: Thurston Analysis

Step								
Step 1: Observation (frequency matrix)	0	1	0	1	3	3	6	4
	22	0	9	11	13	14	11	15
	23	14	0	15	19	15	14	18
	22	12	8	0	13	12	12	14
	20	10	4	10	0	9	10	12
	20	9	8	11	14	0	13	15
	17	12	9	11	13	10	0	14
	19	8	5	9	11	8	9	0
Step2: Proportion matrix (obs./N)	0	0.0435	0	0.0435	0.1304	0.1304	0.2609	0.1739
	0.9565	0	0.3913	0.4783	0.5652	0.6087	0.4783	0.6522
	1	0.6087	0	0.6522	0.8261	0.6522	0.6087	0.7826
	0.9565	0.5217	0.3478	0	0.5652	0.5217	0.5217	0.6087
	0.8696	0.4348	0.1739	0.4348	0	0.3913	0.4348	0.5217
	0.8696	0.3913	0.3478	0.4783	0.6087	0	0.5652	0.6522
	0.7391	0.5217	0.3913	0.4783	0.5652	0.4348	0	0.6087
	0.8261	0.3478	0.2174	0.3913	0.4783	0.3478	0.3913	0
Step3: Scale difference (Z-scores)	-	-1.7117	-	-1.7117	-1.1243	-1.1243	-0.6407	-0.9388
	1.7117	-	-0.2759	-0.0545	0.1642	0.2759	-0.0545	0.3912
	-	0.2759	-	0.3912	0.9388	0.3912	0.2759	0.781
	1.7117	0.0545	-0.3912	-	0.1642	0.0545	0.0545	0.2759
	1.1243	-0.1642	-0.9388	-0.1642	-	-0.2759	-0.1642	0.0545
	1.1243	-0.2759	-0.3912	-0.0545	0.2759	-	0.1642	0.3912
	0.6407	0.0545	-0.2759	-0.0545	0.1642	-0.1642	-	0.2759
	0.9388	-0.3912	-0.781	-0.2759	-0.0545	-0.3912	-0.2759	-
Step4: Interval scale estimate	0.9064	-0.2697	-0.3817	-0.2405	0.0660	-0.1542	-0.0800	0.1538
	(column Average)							

Montag's Confidence Estimation

The confidence interval of 95% (95% CI) are calculated that define a range in which if the experiment is repeated and results are recalculated, the population estimate can be found 95% of the time. The observed standard deviation (empirical data fit) is estimated by using the following equation:

$\sigma = b1(n - b2)b3\ (N - b4)b5$ … (Eqn 3).

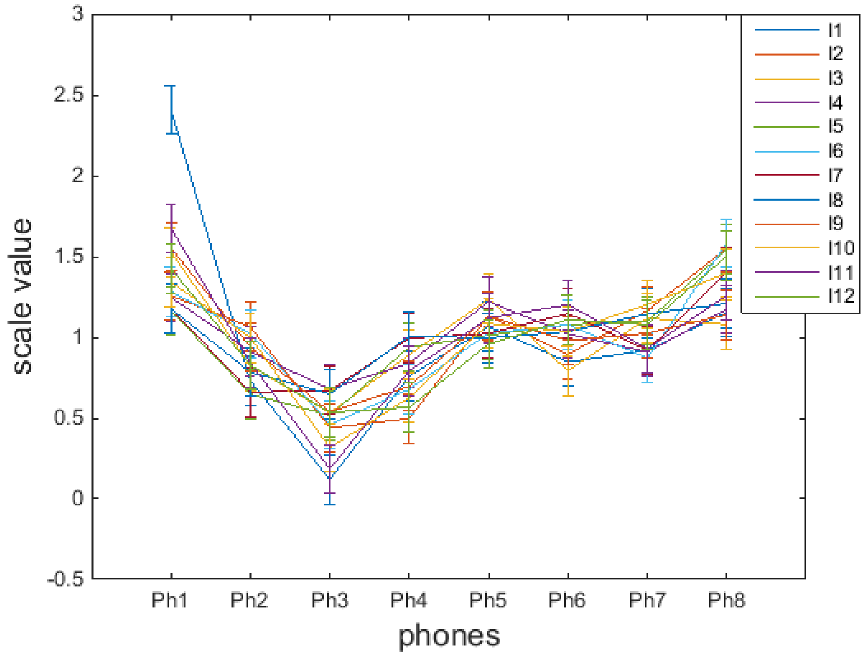
$b1$=1.76, $b2$=3.08, $b3$=-0.613, $b4$ = 2.55, $b5$= -0.491

N is number of observers, n is number of stimuli

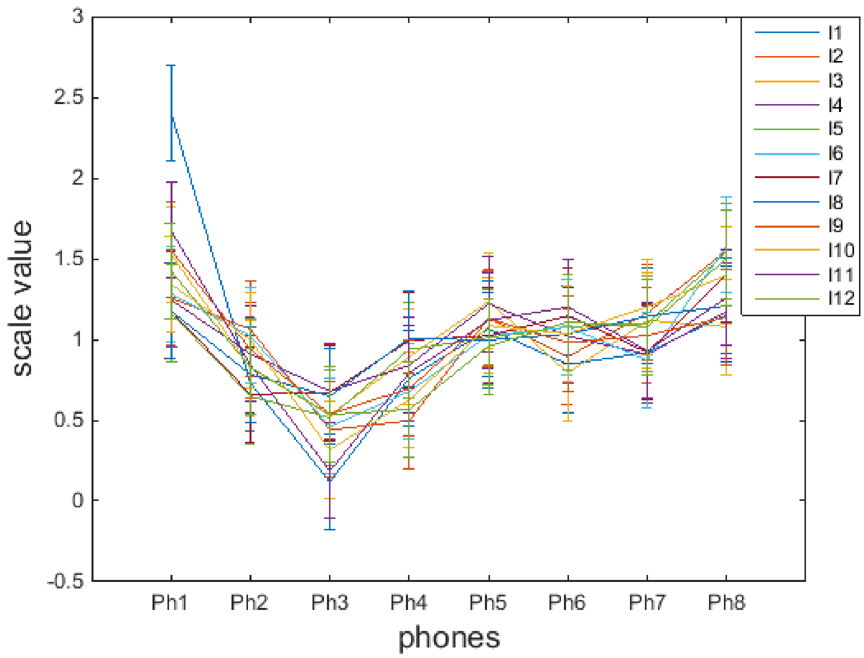
Confidence interval by following:

95% CI=1.96 $\sigma\sqrt{N}$ ... (Eqn 4).

The confidence interval was calculated using the Montag method [29]; for press 1, this value is 0.0356. The errors bars for the CI are smaller and consistent relative to the error bars for the standard deviation. The standard deviation shows the variation among observers for each image set and thus the variance for each image set can be calculated. The variance for press 1 is 0.008, implying that the observers did not vary much in their preference. The appendix shows the plots for error bars with standard deviation and Confidence Intervals for all the presses.



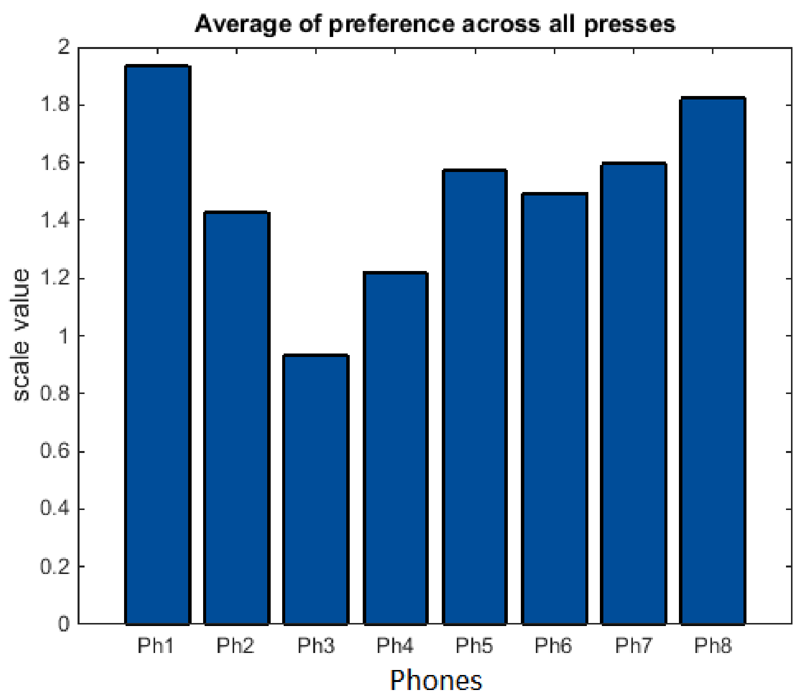
Errors bars with Standard Deviation



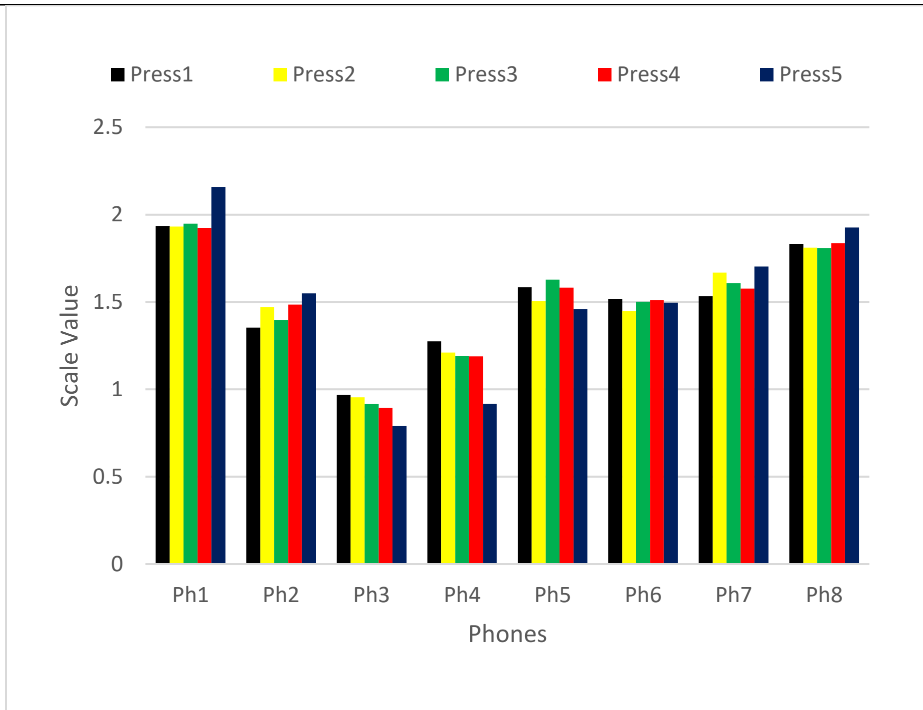
Error bars with Confidence Interval

Figure 4.2-1: Standard deviation and confidence interval for all phones

Preference of each phones across all presses shows the performance of each phone. Figure 4.2-2 shows the preference plot of average performance of phones on different presses. It is observed that preference ratings are similar irrespective of the presses that printed them. This means the image quality from the phones translates well to printed image quality and there is minimal quality loss during the printing process. Phones have bigger impact on the image quality compared to presses as the preference for each phone is similar irrespective of the presses they were printed on. Ph1 and Ph8 have high preference scale, whereas Ph3 has the lowest preference across the board.



Preference plot averaged across the presses



Preference plot of phones across each individual press.

Figure 4.2-2: Average performance of each phone across each press.

Image enhancement was used on one of the presses as a parameter for printing. All the images were treated in the Raster Image Processing (RIP) to improve contrast, sharpness, color tone, and to reduce noise and uneven tones before printing. Figure 4.2-3 shows the preferences for the press with and without image enhancement. Some of the enhanced images were highly preferred but others were affected negatively by the enhancement, making their reception poor. Images from Ph4 were not received well after enhancement whereas images from Ph1 received better ratings. One of the contributing factors in this is that only 11 observers performed the experiment with the Enhanced image set whereas 24

observers performed the default image set experiment. Relatively no change is observed due to enhancement in other phones.

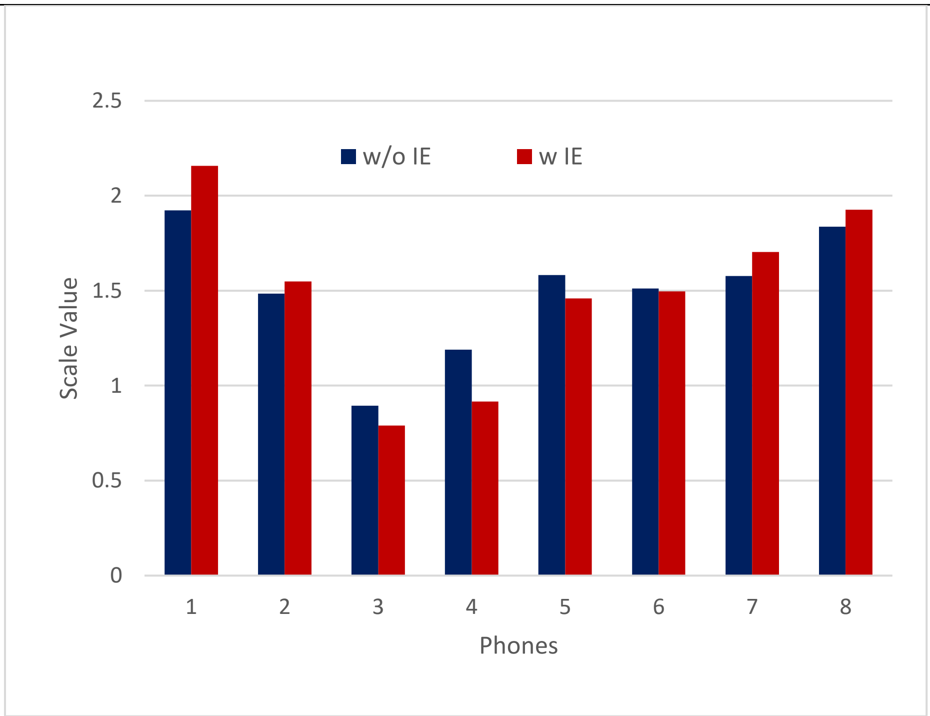
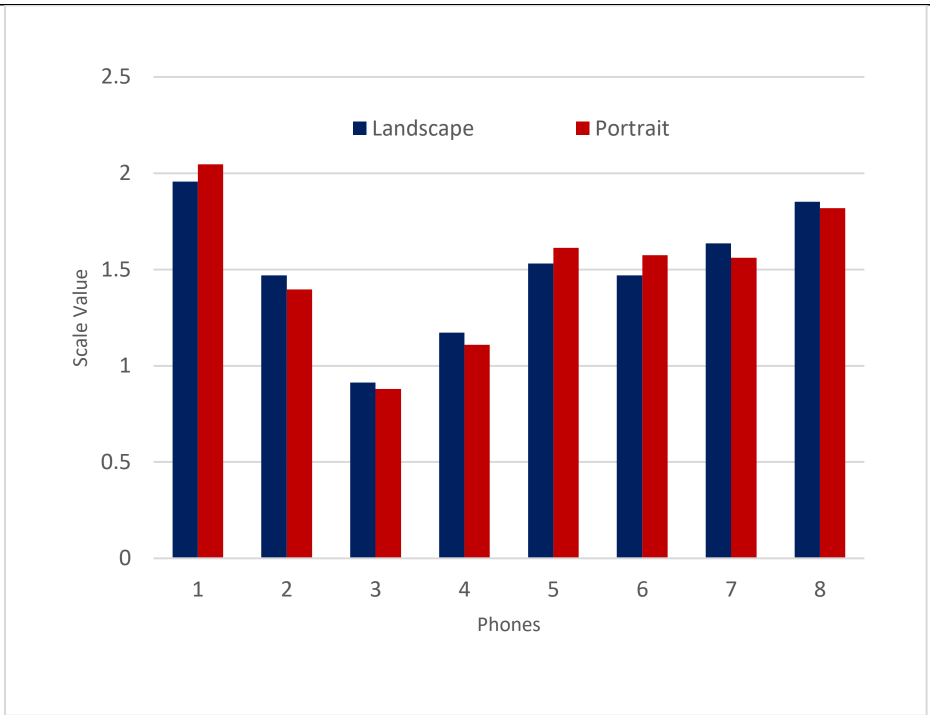
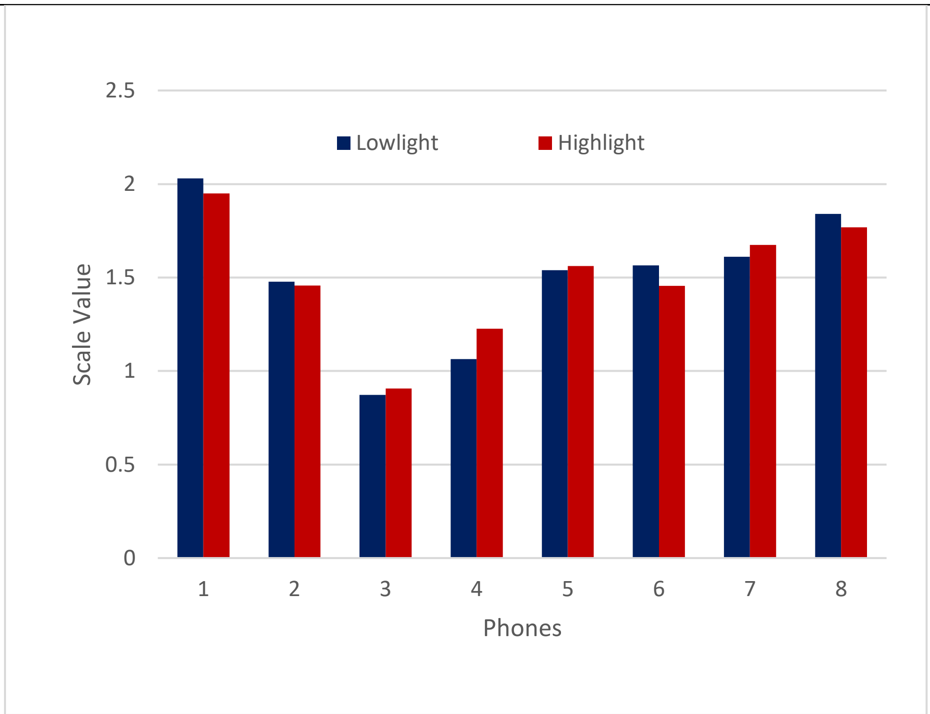


Figure 4.2-3: Preference change due to image enhancement

The study involved three portrait-oriented pictures and nine landscape-oriented pictures. In comparing the average scale values of landscape pictures to the average portrait pictures, it can be seen that the orientation of the pictures does not impact the preference scores. The bar plots of both the modes shown in Figure 4.2-4 validate the average bar plot that was observed for the overall experiment shown in Figure 4.2-2. In both cases, Ph1 and Ph8 are still the most preferred and Ph3 and Ph4 are still the least preferred.



Average of landscape images against portrait images



Average of Low light images against Highly illuminated images

Figure 4.2-4: Preference plots of landscape and portrait images

The performance of smartphones images by scene light level is of particular interest since this lower light levels can prove challenging [29]. Low light levels make the images grainy and noisy. Newer smartphones (from 2015) try to compensate by processing the image to improve the image quality. Ph1, Ph6, Ph8 generally captured the low light levels scenes better than other phones, which is likely a result of better low light performance algorithms. The preference rating improves for phones 6, 7 and 8 as they capture low light levels images better than other phones. Phones 3 and 4 suffer under low light thus their preference rating is lower for low light images. The second bar plot in Figure 4.2-4 shows the comparison between averages of low light levels compared to images with high illumination.

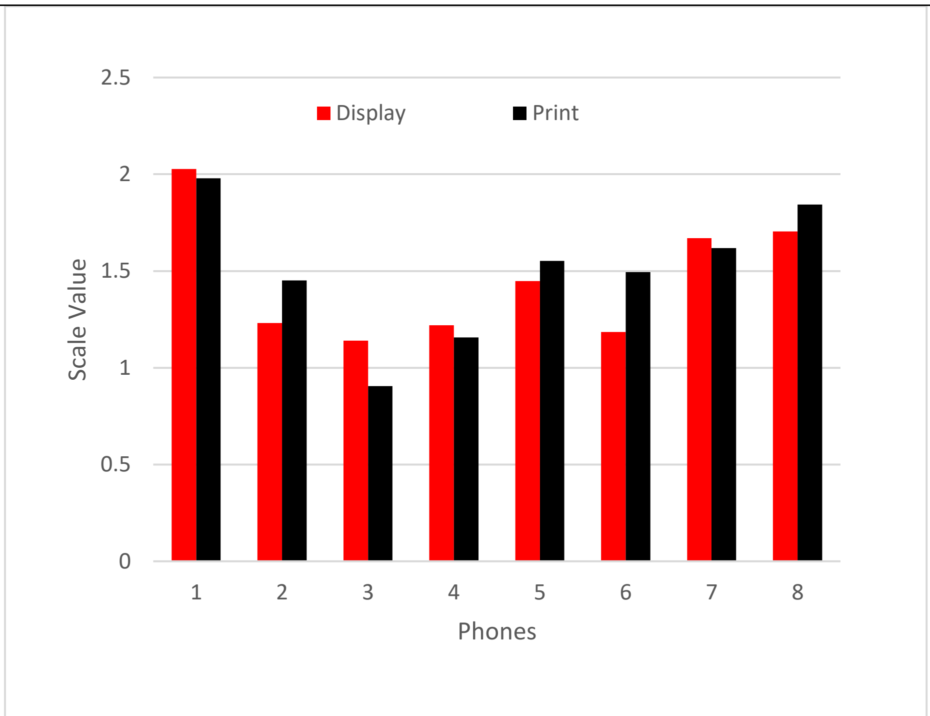


Figure 4.2-5: Average ratings of displayed and printed images

The transition to printed images from displayed images retains the image quality. Ph3 performs the worst in both cases but when images are displayed the difference of scale value is close to Ph2 and Ph3 whereas using printed samples Ph3 has the lowest values. Mostly displayed samples have very similar preferences to printed samples but as it is an indirect scale the preference for display is not necessarily higher than print. There is a high correlation between displayed and printed samples with a correlation coefficient of 0.86. The important thing to note is that the best preferred phone images in display remained the best preferred in print and vice versa. For phones 2 & 6 there was a much higher preference for printed samples as compared to displayed samples. Ph2, Ph6 and Ph8 were good at capturing low light images which create better highlights and shadows in the prints. Thus, printed samples for these phones are more preferred. Ph3 had least amount of noise reduction, which translated to grainy printed samples, thus lowering their preference. Also, Ph3 had lower resolution as compared to other phones, which pixelates the displayed images making them undesirable.

5. Summary & Conclusion

The increasing number of images captured by cell phones generates the necessity to measure the image quality of cell phones. A number of these images are converted to printed samples. Digital printing is a preferred choice as these presses produce good image quality and are best suited for variable data printing that goes hand in hand with the low volume prints that individual cell phones generate.

Image experts have developed a keen sense to fish out defects over the course of their careers. A stark difference is observed between the preferences of image experts and non-image experts. The first experiment highlights the need for having a healthy balance of image experts and non-image experts to get proper results that represent the real-world image quality preference. The second experiment concentrates the efforts of image quality preference on to printed samples generated using smart phone cameras. Smartphones have varying capabilities of image capture depending on the optics and image processing algorithms. It is important to figure out if the smart phone quality affects the printed image quality. The data shows that Ph1 and ph8 have the highest preference across all digital presses Ph3 and ph4 have the lowest preference. Ph5, Ph6 and Ph7 have similar preferences which are towards the high preference scale. The image quality of captured images translates with almost one to one correspondence to the printed samples with a correlation coefficient on 0.86. Image enhancement viz. noise reduction, tone correction etc. was used a parameter for one of the presses. Low light images were more preferred after

enhancement but other images lost their natural look decreasing their preference. The smartphones have a higher impact on image quality as compared to the digital presses which affects the image quality observed in the printed samples. Thus, smartphones affect image quality more than the presses themselves.

Image quality is complex phenomenon and is affected by various factors. The results observed showed that the phone with the best camera on paper translated to very good printed images. One of the most commercially successful smartphone cameras having the best specifications on paper, however, did not perform as well as expected. Some of the moderate cell phones performed well in low light scenes but did not fare well in very bright scenarios. One key factor that affected the results was the image processing algorithms that either soften the image or increase its sharpness. It was observed that sharpened images were generally more preferred than the softened images.

For future work, the study can be expanded to use a larger variety of cell phones. Parameters like high dynamic range and saturation could have a effect on the overall image quality and there has been a vast improvement in low light processing which enhances images in very low lighting conditions. It would be of interest to see how these affect the print quality when these images are printed.

6. References

[1] P. Rosa, "ResearchGate," 2015. [Online]. Available: https://www.researchgate.net/figure/8-Schematic-illustration-of-the-lithography-offset-printing-process-Source-Adapter_fig6_299395183.

[2] J. Sagu, "ResearchGate," 2015. [Online]. Available: https://www.researchgate.net/figure/The-flexographic-printing-process_fig1_275407192.

[3] P. Littlewood, "How it works: Digital Printing," 2015. [Online]. Available: http://fulprint.com/digital-print/digital-printing/#.

[4] W. E. Kasdorf, The Columbia Guide to Digital Publishing, Columbia University Press, 2003.

[5] H. A. o. G. R. Centres, 2015. [Online]. Available: https://phys.org/news/2015-05-inkjet-kesterite-solar-cells.html.

[6] DP3, "Electrophotographic," [Online]. Available: http://www.dp3project.org/technologies/digital-printing/electrophotographic.

[7] R. Berns, "R. S. Berns, Billmeyer and Saltzman's Principles of Color," John Wiley, New York, 2000.

[8] J. Albers, Interaction of Color, Yale university press, 2013.

[9] "Color Mixing," [Online]. Available: https://www.d.umn.edu/~mharvey/th1501color.html.

[10] K. Beaty, "Printing Lingo: What is 4-Color Process Printing?," 2018. [Online]. Available: https://www.formaxprinting.com/blog/2018/09/printing-lingo-what-is-4-color-process-printing/.

[11] L. Candela, "5.2 Raster Image Processing," [Online]. Available: https://courses.lumenlearning.com/suny-graphicdesign/chapter/5-2-raster-image-processing/.

[12] A. T. Young, "Digital Halftoning," 2012. [Online]. Available: https://aty.sdsu.edu/bibliog/latex/scan/halftone.html.

[13] W. E. Kasdorf, "The columbia guide to digital publishing," in *The columbia guide to digital publishing*, 2003.

[14] M. E.-N. T. M. K. Marianne Hörlesberger, Challenges in the Management of New Technologies, World Scientific, 2007.

[15] R. Chung and M. Rees, A Survey of digital and offset print quality issues, 2007.

[16] G. A. Gescheider, Psychophysics: The Fundamentals, Laurance Earlbaum Associates, 1997.

[17] T. Engen, Psychophysics. In: States of Brain and Mind, Birkhäuser, 1988.

[18] H.-C. Lee, Introduction to Color Imaging Science, Cambridge University Press, 2005.

[19] L. L. Thurstone, Rank Order as a Psychophysical Method, University of Chicage, 1931.

[20] P. G. Engeldrum, Psycometric Scaling, Imcotek Press, 2000.

[21] Samsung. [Online]. Available: https://www.samsung.com/global/galaxy/galaxy-s7/.

[22] HTC. [Online]. Available: https://www.htc.com/us/go/buy-now-htc-one-m8/.

[23] Apple. [Online]. Available: https://www.apple.com/iphone/.

[24] LG. [Online]. Available: https://www.lg.com/us/cell-phones/lg-D800-g2.

[25] Nokia. [Online]. Available: https://www.nokia.com/.

[26] Sony. [Online]. Available: https://www.sony.com/.

[27] Xerox. [Online]. Available: https://www.xerox.com/.

[28] Shutterfly. [Online]. Available: https://www.shutterfly.com/.

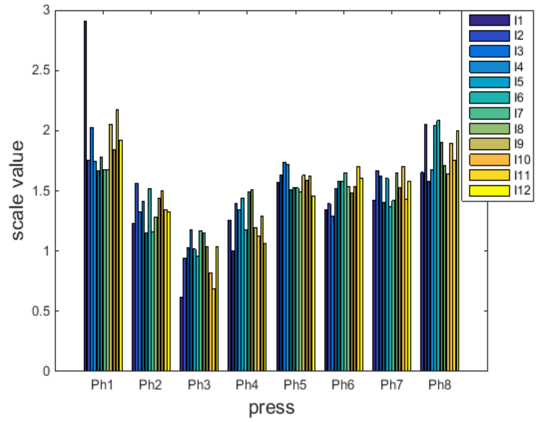
[29] E. W. Jin, J. B. Phillips, S. Farnand, M. Belska, V. Tran, E. Chang, Y. Wang and B. Tseng, "Towards the Development of the IEEE P1858 CPIQ Standard – A validation study," *Electronic Imaging,* 2017.

[30] P. G. Engeldrum, Psychometric Scaling: A Toolkit for Imaging Systems., Massachusetts: Imcotek Press, 2000..

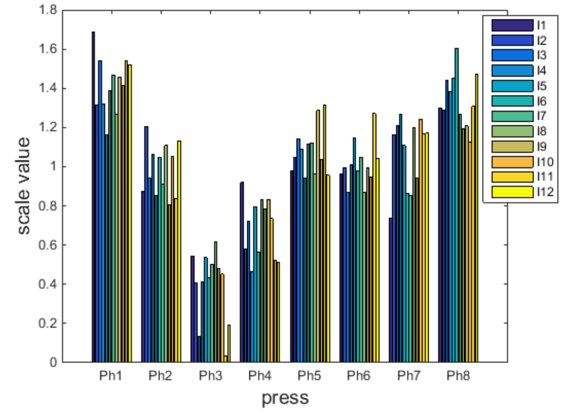
[31] E. W. &. K. B. W. Jin, "Slider-adjusted softcopy ruler for calibrated image quality assessment.," no. Journal of Electronic Imaging, 19(1), 011009-011009, 2010..

[32] E. W. Jin, J. B. Phillips, S. Farnand, M. Belska, V. Tran, E. Chang, Y. Wang and B. Tseng, *Towards the Development of the IEEE P1858 CPIQ Standard – A validation study,* Electronic Imaging, Image Quality and System Performance XIV, pp. 88-94(7).

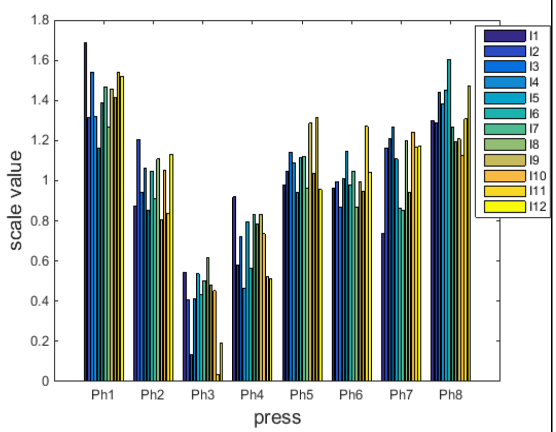
7. Appendix



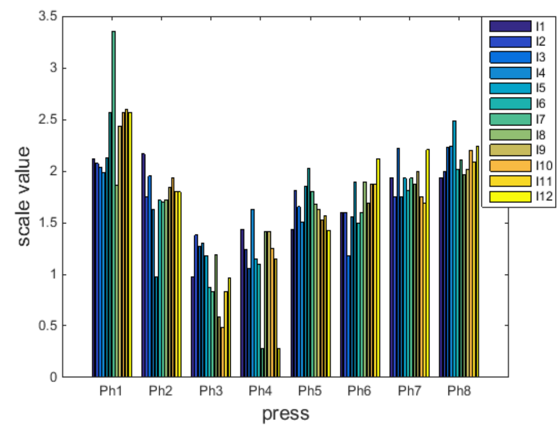
Press 1



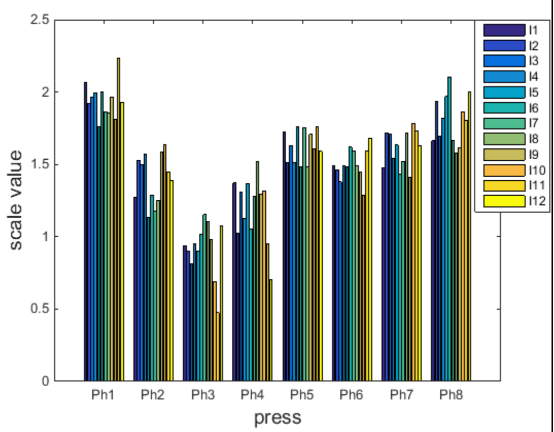
Press 2



Press 3

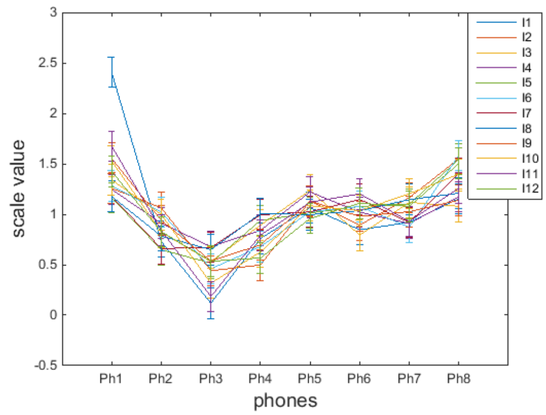


Press 4

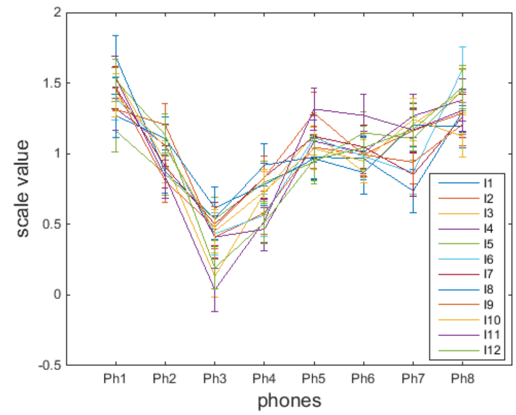


Press 5

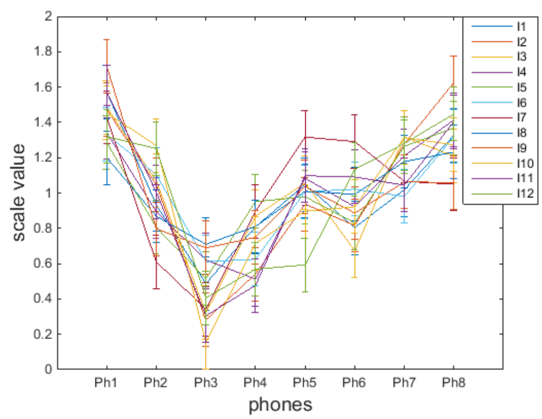
Figure 7-1: Preference plots of 5 presses.



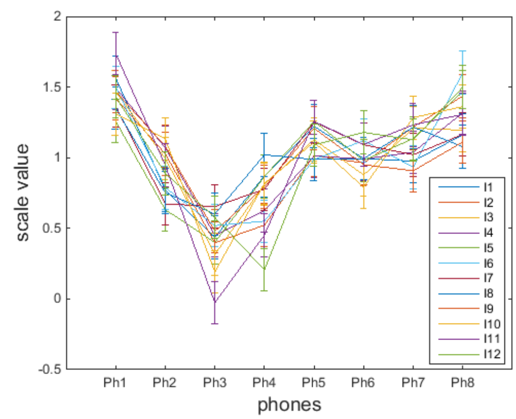
Press 1



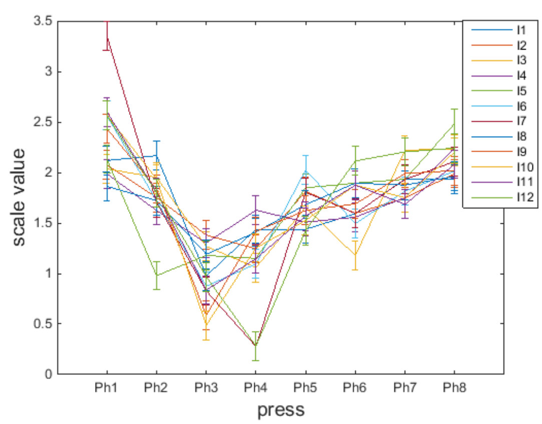
Press 2



Press 3

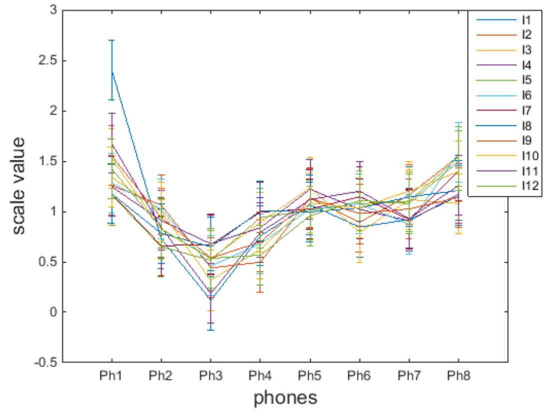


Press 4

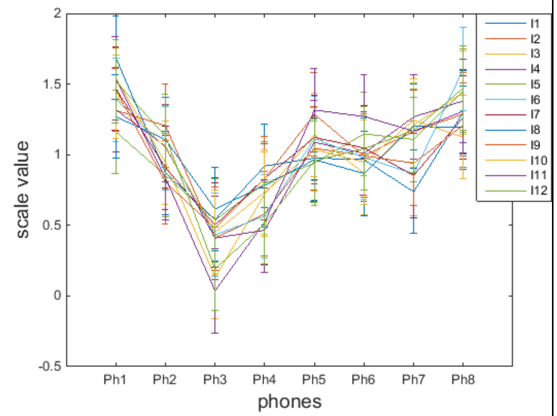


Press 5

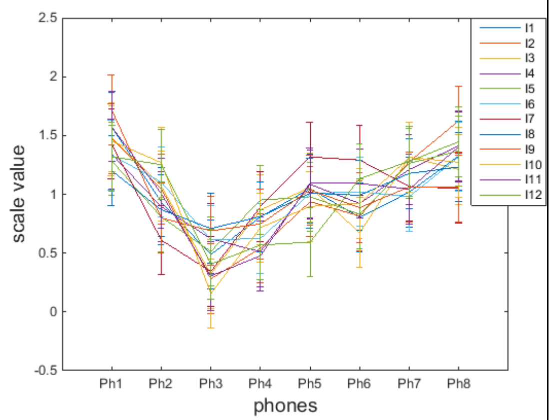
Figure 7-2: Standard deviation error bars of 5 presses.



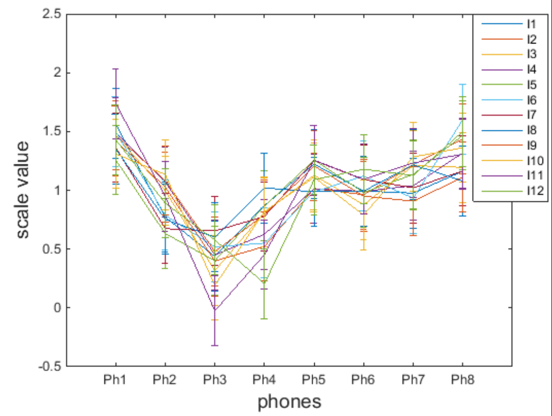
Press 1



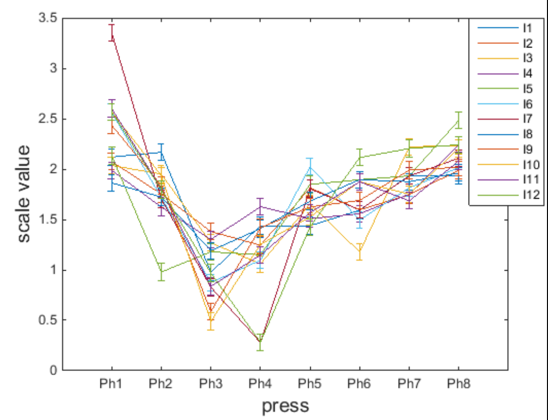
Press 2



Press 3

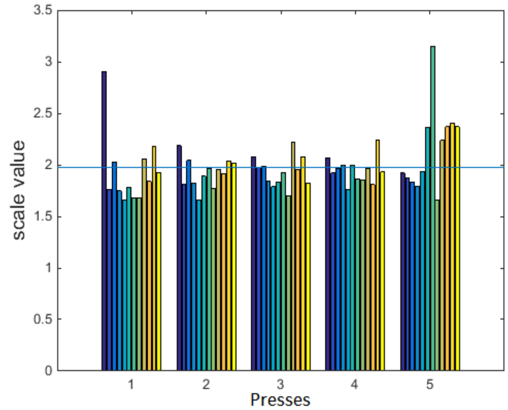


Press 4

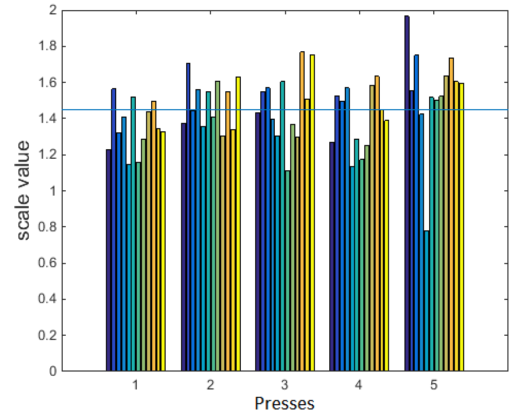


Press 5

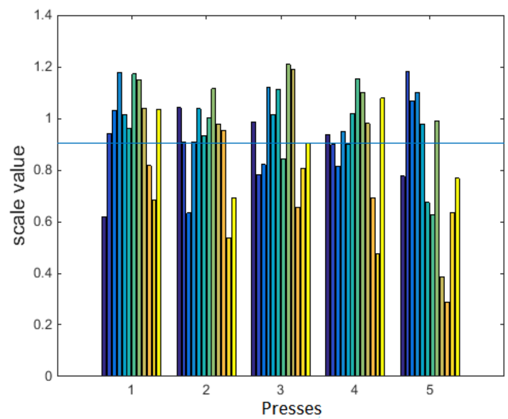
Figure 7-3: Confidence Interval error bars of 5 presses.



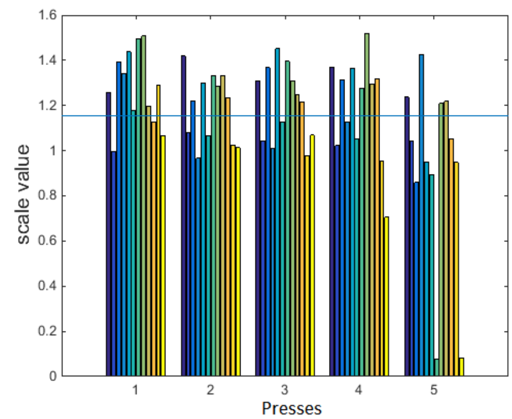
Phone 1



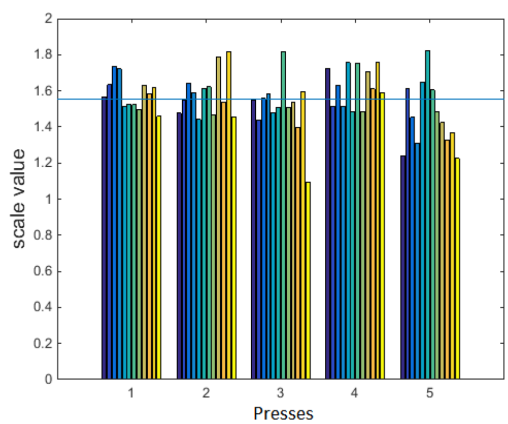
Phone 2



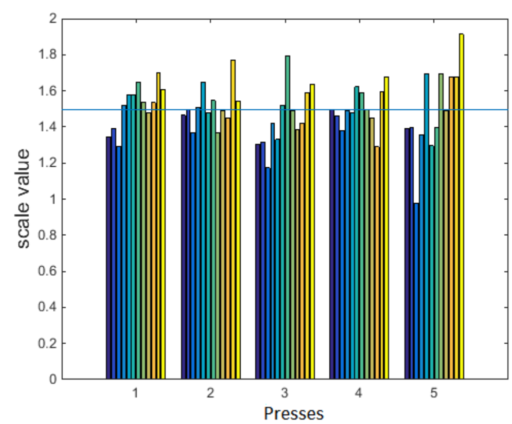
Phone 3



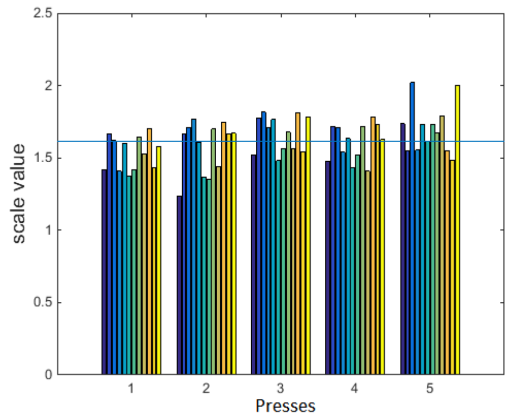
Phone 4



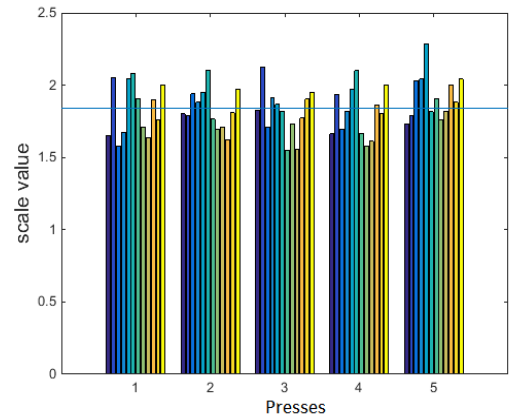
Phone 5



Phone 6



Phone 7



Phone 8

Figure 7-4: Preference plots of each phone across five presses.