National Technical University of Athens
School of Electrical and Computer Engineering
Division of Communication, Electronics and Information Engineering

Development of a smartphone-based sensor for urine protein measurement

Diploma Thesis

of

MARIA TSIRONI

Supervisor: Evangelos Hristoforou
Professor, NTUA

Athens, July 2024

National Technical University of Athens
School of Electrical and Computer Engineering
Division of Communication, Electronics and Information Engineering

Development of a smartphone-based sensor for urine protein measurement

Diploma Thesis

of

MARIA TSIRONI

Supervisor: Evangelos Hristoforou
Professor, NTUA

Approved by the examination committee on 5th July 2024.

(Signature)	*(Signature)*	*(Signature)*
.....	.....	.....
Evangelos Hristoforou	George Matsopoulos	Konstantinos Politopoulos
Professor, NTUA	Professor, NTUA	Associate Professor, NTUA

Athens, July 2024

National Technical University of Athens
School of Electrical and Computer Engineering
Division of Communication, Electronics and Information Engineering

DISCLAIMER ON ACADEMIC ETHICS AND INTELLECTUAL PROPERTY RIGHTS

Being fully aware of the implications of copyright laws, I expressly state that this diploma thesis, as well as the electronic files and source codes developed or modified in the course of this thesis, are solely the product of my personal work and do not infringe any rights of intellectual property, personality and personal data of third parties, do not contain work / contributions of third parties for which the permission of the authors / beneficiaries is required and are not a product of partial or complete plagiarism, while the sources used are limited to the bibliographic references only and meet the rules of scientific citing. The points where I have used ideas, text, files and / or sources of other authors are clearly mentioned in the text with the appropriate citation and the relevant complete reference is included in the bibliographic references section. I fully, individually and personally undertake all legal and administrative consequences that may arise in the event that it is proven, in the course of time, that this thesis or part of it does not belong to me because it is a product of plagiarism.

(Signature)

.....
Maria Tsironi
Graduate of Electrical
and Computer
Engineering, NTUA
5th July 2024

Abstract

Proteinuria is the presence of proteins in urine above a certain threshold. Because proteinuria is a common finding in primary care practice, there exists a large body of research regarding the underlying conditions indicated by different types of proteinuria, the pathological thresholds and the measurement techniques that provide the most accurate results. This thesis is focused toward finding a novel, inexpensive method for proteinuria detection in home or point-of-care testing. The method proposed is based on smartphone-acquired image colorimetry where the color of a urine sample mixed with the Bradford reagent, an assay that changes color upon binding to protein, is evaluated. A test platform is also developed for the calibration of illumination conditions and a project-oriented coding implementation is described in OpenCV.

Keywords

Proteinuria, protein measurement, Bradford assay, spectrophotometry, digital image colorimetry

Περίληψη

Πρωτεϊνουρία ονομάζεται η παρουσία πρωτεϊνών στα ούρα πάνω από ένα συγκεκριμένο όριο. Καθώς η πρωτεϊνουρία αποτελεί συχνό εύρημα στην πρακτική της πρωτοβάθμιας περίθαλψης, υπάρχει μεγάλος όγκος έρευνας σχετικά με τις υποκείμενες παθήσεις που υποδηλώνουν οι διάφοροι τύποι πρωτεϊνουρίας, τα παθολογικά όρια και τις τεχνικές μέτρησης που παρέχουν τα πιο ακριβή αποτελέσματα. Η παρούσα εργασία επικεντρώνεται στην εξεύρεση μιας νέας, ανέξοδης μεθόδου για την ανίχνευση της πρωτεϊνουρίας κατ' οίκον ή σε σημεία εξέτασης. Η μέθοδος που προτείνεται βασίζεται σε χρωματομετρία εικόνας που λαμβάνεται με κάμερα κινητού, όπου αξιολογείται το χρώμα ενός δείγματος ούρων που αναμιγνύεται με το αντιδραστήριο Bradford, μια ουσία που αλλάζει χρώμα κατά τη δέσμευση με πρωτεΐνη. Κατασκευάστηκε επίσης μια πλατφόρμα μέτρησης για την εξισορρόπηση των συνθηκών φωτισμού, ενώ περιγράφεται η ανάπτυξη κώδικα με χρήση της βιβλιοθήκης OpenCV που μπορεί να εξάγει χρωματικές τιμές από τις σχετικές εικόνες.

Λέξεις Κλειδιά

Πρωτεϊνουρία, μέτρηση πρωτεϊνών, αντιδραστήριο Bradford, φασματοφωτομετρία, ψηφιακή χρωματομετρία εικόνας

Acknowledgements

I would like to thank Professor Hristoforou for the assignment of this project and his guidance during these months. It has been a pleasure learning from you in the classroom and in the lab. I would also like to thank Mr. Angelo Ferraro for his constant supervision and unwavering support throughout the process. I am sure a lot of patience was required to teach me chemistry procedures 101.

I really appreciate Andreas' moral support, practical help and cooking skills, as well as my old and new friends that listened to my venting and nagging all this year. An honorable mention to my brother for making our home more crowded, but life definitely more interesting.

Athens, July 2024

Maria Tsironi

Table of Contents

List of Figures

List of Tables

Εκτεταμένη περίληψη στα Ελληνικά

Η αυξημένη παρουσία πρωτεϊνών στα ούρα είναι συχνά δείκτης διάφορων παθήσεων, όπως η χρόνια νεφρική ανεπάρκεια, ο διαβήτης ή οι καρδιαγγειακές παθήσεις. Για τον λόγο αυτό έχει γίνει εκτενής έρευνα σχετικά με τους διαφορετικούς τύπους πρωτεϊνουρίας και τις σχετιζόμενες παθήσεις, αλλά και τον τρόπο με τον οποίο γίνεται η μέτρηση των πρωτεϊνών στα ούρα ώστε να προκύψουν αξιόπιστα αποτελέσματα. Πέρα από τις ακριβείς αλλά χρονοβόρες εργαστηριακές εξετάσεις, η πιο συνηθισμένη προκαταρκτική μέθοδος για μία εκτίμηση της συγκέντρωσης του δείγματος σε πρωτεΐνες είναι η ταινία εξέτασης ούρων (dipstick) η οποία μπορεί να περιέχει και άλλα αντιδραστήρια πέρα από αυτό για την ανίχνευση πρωτεϊνών. Έχει αποδειχθεί ωστόσο ότι η συγκεκριμένη μέθοδος, παρά τη χρηστικότητά της, δεν προσφέρει σε πολλές περιπτώσεις αξιόπιστες μετρήσεις, με αποτέλεσμα να γίνονται προσπάθειες για την ανάπτυξη εναλλακτικών τεχνικών.

Στην έρευνα που παρουσιάζεται εδώ χρησιμοποιήθηκε για την ανίχνευση των πρωτεϊνών το αντιδραστήριο Bradford, το οποίο εμφανίζει κατά τη σύνδεσή του με μόρια μεγάλη επιλεκτικότητα στις πρωτεΐνες. Το διάλυμα Bradford περιέχει συγκεκριμένα το αντιδραστήριο Coomassie Brilliant Blue G-250, το οποίο έχει κόκκινο χρώμα που μετατρέπεται σε μπλε μετά τη σύνδεσή του με πρωτεΐνη. Ουσιαστικά στο Bradford η αντιδρώσα ουσία περιέχεται σε τρεις συνυπάρχουσες διαφορετικές μορφές, την ανιονική μορφή μπλε χρώματος με μέγιστο απορρόφησης στα 595 nm, την ουδέτερη μορφή πράσινου χρώματος με μέγιστο απορρόφησης στα 650 nm, η οποία έχει προκύψει αφού έχει δεχθεί ένα πρωτόνιο, και η κατιονική μορφή κόκκινου χρώματος με μέγιστο απορρόφησης στα 470 nm, η οποία έχει προκύψει αφού έχει δεχθεί δύο πρωτόνια.

Στα πειράματα που πραγματοποιήθηκαν με το φασματοφωτόμετρο, η μέτρηση της απορρόφησης έγινε στα 595 nm. Με σκοπό τον προσδιορισμό μιας καμπύλης βαθμονόμησης του αισθητήρα, μετρήθηκαν διαφορετικές συγκεντρώσεις δειγμάτων με διαφορετικές συγκεντρώσεις του διαλύματος Bradford σε διάφορες αναλογίες. Επιπλέον, έγινε και μία σύγκριση των αποτελεσμάτων ανάμεσα σε μία καμπύλη με υδατικά δείγματα πρωτεϊνών και σε μία καμπύλη που προέκυψε από την προσθήκη αυξανόμενων

ποσοτήτων πρωτεΐνης σε δείγμα ούρων. Βασικό ζητούμενο στην καμπύλη ήταν η όσο το δυνατόν μεγαλύτερη γραμμικότητα σε ένα ευρύ φάσμα συγκεντρώσεων που περιλαμβάνει τις οριακές τιμές για την ανίχνευση ή μη πρωτεϊνουρίας.

Στη συνέχεια, εξετάστηκε το κατά πόσο η εξαγωγή του χρώματος από φωτογραφία των δειγμάτων ανεμειγμένων με το διάλυμα Bradford μπορεί να προσομοιάσει την καμπύλη βαθμονόμησης του φασματοφωτόμετρου. Για τα συγκεκριμένα πειράματα χρησιμοποιήθηκαν 48- ή 96-well plates στα οποία τοποθετήθηκαν διαλύματα διαφορετικών συγκεντρώσεων. Διερευνήθηκε επίσης και το αν μπορεί να ληφθεί αξιόπιστη μέτρηση από την προσθήκη του διαλύματος Bradford πάνω σε χαρτί, όπου διαπιστώθηκε έντονο coffee ring effect και επίδραση της κυτταρίνης του χαρτιού στη μέτρηση, πρόβλημα το οποίο λύθηκε με επιλογή χαρτιού από γυάλινες μικροΐνες.

Η επεξεργασία των εικόνων έγινε σε πρωταρχικό στάδιο στο πρόγραμμα ImageJ με το οποίο μετρήθηκαν οι μεταβλητές των χρωματικών μοντέλων RGB, HSV και L*a*b*. Το κανάλι b* του μοντέλου L*a*b* επιλέχθηκε ως αυτό που δίνει αποτελέσματα πλησιέστερα στην καμπύλη βαθμονόμησης. Πραγματοποιήθηκαν επίσης κάποια επαναληπτικά πειράματα με σκοπό τη μείωση του συνολικού όγκου των εξεταζόμενων δειγμάτων και τη βελτίωση της απόκρισης σε αυτές τις συνθήκες φωτισμού.

Ως προτεινόμενη διάταξη για την τελική μέτρηση σχεδιάστηκε και εκτυπώθηκε μία πλατφόρμα μετρήσεων η οποία έχει μικρές διαστάσεις και περιλαμβάνει τρία "δοχεία" ή "πηγάδια", σε καθένα από τα οποία περιέχεται κάποιο διάλυμα με την ίδια περιεκτικότητα σε Bradford: ένα δείγμα υπό μέτρηση, ένα με μηδενική συγκέντρωση και ένα με τη μέγιστη ανιχνεύσιμη συγκέντρωση. Λόγω της κανονικοποίησης του χρώματος αλλά και της απορρόφησης του φασματοφωτόμετρου σε τιμές από 0 μέχρι 1, τα αποτελέσματα από τη μέτρηση αυτή μπορούν να συσχετιστούν ημι-ποσοτικά με τη συγκέντρωση του υπό μέτρηση δείγματος.

Σε τελικό στάδιο αναπτύχθηκε επίσης κώδικας με χρήση της βιβλιοθήκης OpenCV σε python για την αυτοματοποίηση της διαδικασίας. Η υλοποίηση ανιχνεύει τους κύκλους στην εικόνα και εξάγει τον μέσο όρο της b* τιμής εντός κάθε κύκλου, μετά την εφαρμογή φίλτρου για τη μείωση του θορύβου, των σκιάσεων και των αντανακλάσεων.

Στο τέλος της εργασίας συζητούνται τα συνολικά αποτελέσματα καθώς και προτάσεις για μελλοντική έρευνα, όπως η ανάγκη για περισσότερη μελέτη με διαλύματα συνθετικών ούρων και η χρήση διαφορετικών πρωτεϊνών για την καμπύλη βαθμονόμησης, ενώ επισημαίνονται και πιθανές βελτιώσεις της προτεινόμενης μεθόδου.

Chapter 1

Introduction

Proteinuria, the presence of proteins in urine, is a common finding in primary care practice. On one hand, the causes of proteinuria can be benign, such as dehydration, intense exercise or fever [8]. On the other hand, persistent proteinuria might indicate more serious cardiovascular and kidney disorders and is an important metric for assessing the progression of chronic kidney disease (CKD) [9]. It is significant then to develop a protocol for the approach of patients with detected proteinuria. In the protocol presented in Kashif et al. [10], the evaluation of proteinuria begins with a positive dipstick test, which is the most popular testing method for proteinuria. In case of detected proteinuria, the dipstick test is repeated and in the case of a second positive result, additional testing has to be performed, such as 24-hour sample evaluation or estimation of the spot urine protein-to-creatinine ratio.

Even though the dipstick method is preferred for its low cost and convenience, its sensitivity is quite low, yielding false positive or false negative results in various cases. Other techniques for the detection of proteinuria include turbidometric tests, which are often cumbersome and lack precision, and electrophoresis, which requires more sophisticated equipment and is exclusively a lab procedure [9]. Therefore, there has been research for the development of a semi-quantitative technique that determines the urine protein with higher accuracy. A plethora of such methods has been based on the existence of highly sensitive colorimetric assays and the colorimetry potential of smartphones. An attempt at such a solution is thus presented here, employing the Bradford assay, where the binding of dye to protein causes a shift in the absorption maximum of the solution, which can be monitored with the proper equipment [11].

In the literature survey section, the theoretical basis of the overall subject is established, beginning from the basics of the protein structure and a description of the renal function to a more detailed overview of the measurement of proteinuria and the existing

solutions. In the experiment preparation section, the materials and the different experiments that were performed are presented, reporting the development of the calibration curve, the various measurements (with the spectrophotometer, on paper, in well plates via image) and the design of the final test platform and the method proposed. In the discussion section the overall results are discussed and the limitations of the method are presented, compared with the existing research on the subject.

Supplementary content is provided in the appendices at the end of the thesis. In Appendix A a more detailed analysis of the image processing is demonstrated, where additional figures are presented and the specifics of the image handling with OpenCV are demonstrated as well. In Appendix B more theoretical concepts are analyzed on the subject of digital image processing, as well as an evaluation of the subject of colorimetry and the development of color spaces.

Chapter 2

Literature Survey

2.1 Protein structure

Proteins are the most intricate and functionally advanced molecules known. In this section, the structure and basic functions of a protein are presented, following the book of Alberts, Johnson, Lewis et al. [1].

A protein molecule consists of a *polypeptide backbone* and the attached *side chains*. In general, the protein is a single chain of amino acids which are linked to each other through a covalent peptide bond. The general formula of an amino acid is $H_2NCHRCOOH$, where H_2N is the amino group, COOH is the carboxyl group and R is one of the 20 commonly found side chains (Table 2.1). The peptide bond between two amino acids is actually formed between the carboxyl group of the first amino acid and the amino group of the second amino acid (shown in Figure 2.1). The polypeptide backbone then is the core of the polypeptide chain consisting of the recurring parts of the amino acids, while the attached side chains that do not participate in the peptide bonds are the components that make proteins distinct. The polypeptide chain has two different ends: one with a free amino group (NH_3^+ or NH_2), called the amino terminus or the N-terminus, and one with a free carboxyl group (COO^- or COOH), called the carboxyl terminus or C-terminus.

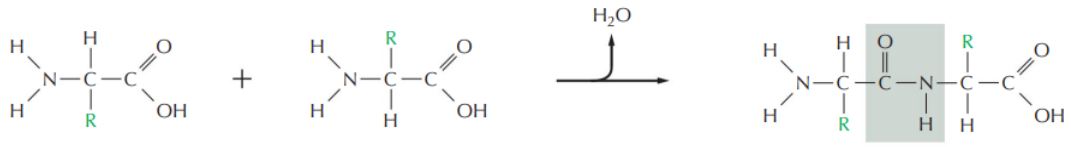


Figure 2.1. *Formation of the peptide bond (gray box) between two amino acids (taken from [1]).*

The folding and thus the shape of a protein is determined by the noncovalent bonds

between different parts of its polypeptide chain, which can be hydrogen bonds, electrostatic attractions and van der Waals attractions. These noncovalent bonds are significantly weaker than covalent bonds but because many of them are simultaneously present in a protein, they can guarantee the stability of the folding.

Amino acid	Abrv	Side chain	Amino acid	Abrv	Side chain
Aspartic acid	Asp	negative	Alanine	Ala	nonpolar
Glutamic acid	Glu	negative	Glycine	Gly	nonpolar
Arginine	Arg	positive	Valine	Val	nonpolar
Lysine	Lys	positive	Leucine	Leu	nonpolar
Histidine	His	positive	Isoleucine	Ile	nonpolar
Asparagine	Asn	uncharged polar	Proline	Pro	nonpolar
Glutamine	Gln	uncharged polar	Phenylalanine	Phe	nonpolar
Serine	Ser	uncharged polar	Methionine	Met	nonpolar
Threonine	Thr	uncharged polar	Tryptophan	Trp	nonpolar
Tyrosine	Tyr	uncharged polar	Cysteine	Cys	nonpolar

Table 2.1. *The 20 most common amino acids in proteins with their 3-letter abbreviations and type of their side chains. The side chains can either be nonpolar and hydrophobic, or polar, negative (acidic), positive (basic) or uncharged [1].*

Besides these noncovalent bonds, another weak force that contributes in the protein folding is the hydrophobic clustering force. This means that in an aqueous environment the nonpolar side chains of a protein, which are hydrophobic, tend to group together in the interior of the molecule to avoid contact with water, while polar side chains tend to arrange themselves on the outside of the molecule, where they can form bonds with water and other polar molecules. The result of these interactions is usually a three-dimensional folded structure called *conformation* that minimizes the protein's free energy.

Most proteins have a length of 50 to 2000 amino acids, with large proteins usually comprised of distinct *protein domains*, structural units folded in some ways independently of each other and that usually contain between 40 and 350 amino acids. Even though the conformation of each protein is unique, two folding patterns are recurring formations in a variety of proteins, the *α helix* and *β sheet*. In the *α* helix "a hydrogen bond forms between every fourth peptide bond, linking the C=O of one peptide bond to the N–H of another" [1], giving the protein a characteristic cylinder shape by forming a helix with a complete turn every 3.6 amino acids. In the *β* sheet, adjacent peptide chains are held together via hydrogen bonds forming a plane (or a sheet), while the amino acid side chains project above and below the plane of the sheet. The *β* sheet might consist either from parallel polypeptide chains (running in the same direction) or

from antiparallel chains (each section of the backbone runs in the opposite direction of its immediate neighboring sections).

In general, proteins can be categorized into *protein families*, considering the resemblances in their amino acid sequence and their three-dimensional conformation. It has also been found that two proteins whose amino acid sequence is identical by more than 25% usually share the same overall structure.

The weak noncovalent bonds that help a protein chain fold into a specific shape also enable proteins to bind together to form larger structures within a cell. Any part of a protein's surface that can interact with another molecule using noncovalent bonds is known as a *binding site*. A protein may have binding sites for different types of large and small molecules. When a binding site recognizes the surface of another protein, the tight connection between the two folded polypeptide chains at this site forms a larger protein molecule with a specific geometric arrangement. Each polypeptide chain within such a protein is called a *protein subunit*. For example, the simplest case is the formation of a *dimer*, where two identical folded polypeptide chains bind to each other in a "head-to-head" configuration.

Concerning the shape of proteins, a characteristic arrangement is that of *globular* proteins, like most enzymes, in which the polypeptide chain folds into a shape resembling a compact ball with an irregular surface. These protein molecules are also able to form long structures if their binding sites are complementary to another site on the surface of the same molecule, often resulting to the creation of a helical structure. Another common shape is that of *fibrous* proteins, where each protein molecule spans a large distance and the overall protein has a fairly simple, elongated three-dimensional configuration. A plethora of fibrous proteins can be especially found outside the cell, where they serve as a primary part of the *extracellular matrix* that aids in uniting groups of cells to form tissues.

The binding of proteins to other molecules is characterized by great specificity, which means that usually a protein molecule can only bind a few molecules comparing to the thousand different ones it encounters. The substance that is bound by the protein, no matter if it is an ion, a small molecule or a macromolecule like another protein, is a *ligand* for that protein. This increased specificity to specific ligands for each protein is achieved with the different surface configurations which lead to unique chemical reactivity. Firstly, the binding sites in a folded polypeptide chain create more solid bonds with ligands when water is kept away. Even though a mechanism that allows

the protein to bind with a specific ligand but not with water might seem improbable, it is possible for a protein to keep a ligand-binding site dry because it is energetically unfavorable for a water molecule to break from its large hydrogen-bonded network and reach into a crevice on the surface of a protein. Moreover, the clustering of neighboring polar amino acid side chains is able to modify their reactivity. For instance, when during a folding many negatively charged side chains are forced together, the affinity of the area for a positively charged ligand is significantly increased, or when neighboring side chains interact with each other through hydrogen bonds, it is possible that typically unreactive groups can become reactive.

Binding between proteins can take place in many ways, such as with a surface-string interaction, where a part of the surface of the first protein binds to an extended loop of a polypeptide chain of the second protein, or with a helix-helix interaction, where two *a*-helices, one from each protein, bind together forming a *coiled coil*, or, most commonly, with a surface-surface interaction, where the rigid surface of one protein exclusively matches the surface of another, often resulting to a very tight, highly selective binding. In the *antibody* protein family, also known as *immunoglobulins*, each antibody is a Y-shaped molecule with two identical *antigen*-binding sites, one on each arm of the Y. These sites "are formed from several loops of polypeptide chain that protrude from the ends of a pair of closely juxtaposed protein domains" [1]. By only altering the length and the amino acid sequence of these loops, a huge diversity of antigen-binding sites can be generated.

Another very important class of proteins is called *enzymes*. Enzymes bind to certain ligands called *test platforms* and rapidly transform them into one or more chemically altered products, speeding up reactions and acting as catalysts that allow cells to control the making and breaking of covalent bonds. Sometimes amino acids alone are not sufficient to perform certain functions, and enzymes, as well as other proteins, have to use small non protein molecules to perform these tasks. Such molecules carried by enzymes might be metal atoms, or small organic molecules that aid their catalytic functions (called *coenzymes*). Coenzymes that cannot be synthesized by humans and therefore have to be provided through diet are called *vitamins*.

2.2 Renal function

In this section the core elements of the urinary system are presented, as found in Ganong's Review of Medical Physiology [2]. The urinary system is comprised of

the *kidneys*, the *bladder* and the *ureters*. Each human kidney has about a million *nephrons*, which are the functional units of the kidney. The main functions of the kidneys include the filtration of the blood plasma, the excretion of metabolic waste, such as urea, ammonium and foreign chemicals, from the body with the formation of urine, the absorption of glucose and amino acids and the regulation of water homeostasis, acid-base homeostasis and blood pressure.

In particular, within the kidneys, a fluid resembling plasma undergoes *glomerular filtration*, that is filtration through the glomerular capillaries into the renal tubules. As this filtrate moves through the tubules, its volume decreases and its composition changes through *tubular reabsorption* (where water and solutes are removed from the tubular fluid) and *tubular secretion* (the secretion of solutes into the tubular fluid), ultimately forming the urine that flows into the renal pelvis. From there, urine passes to the bladder and is released from the body through urination, or *micturition*.

Each nephron consists of a *renal tubule* and its *glomerulus* (Figure 2.2). The glomerulus is a tuft (small network) of capillaries that is enclosed in the expanded end of the nephron, called *Bowman's capsule*. Blood enters the glomerulus via an *afferent arteriole* and leaves via an *efferent arteriole*. The blood in the capillaries is separated from the glomerular filtrate in the Bowman's capsule by two cellular levels: one is the capillary endothelium and the other is the specialized epithelium of the capsule. This filtration network allows all neutral substances with diameter less than 4 nm to freely pass into the Bowman's capsule, while neutral substances with a diameter of more than 8 nm are almost entirely excluded. However, the size of the substances is not the only factor that influences the permeability of the glomerular capillaries. The sialoproteins in the glomerular capillary wall are negative charged, thus repelling the negative charged substances attempting to pass through the wall. The result is that filtration of cationic substances is greater than that of neutral substances, while anionic substances with 4 nm diameter are less than half likely to be filtered compared to neutral substances of the same size. For instance, albumin, a protein with negative charge and effective molecular diameter about 7 nm, "normally has a glomerular concentration only 0.2% of its plasma concentration rather than the higher concentration that would be expected on the basis of diameter alone" [2]. In certain conditions, such as nephritis, the negative charges in the glomerular wall are dispersed, resulting to a higher filtration of albumin and its increased presence in urine, called *albuminuria*.

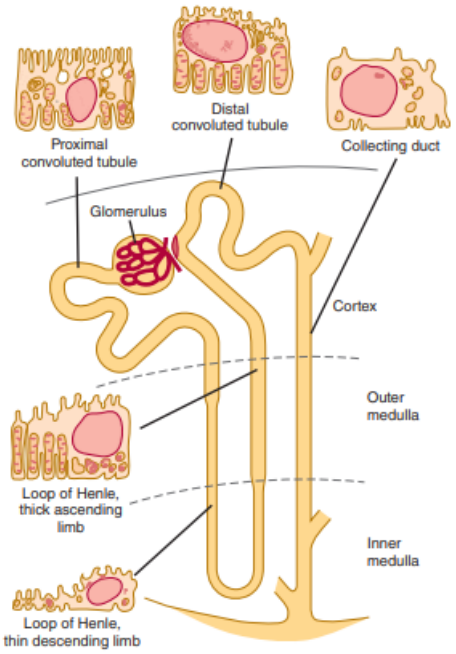


Figure 2.2. *Diagram of a nephron (taken from [2]).*

The glomerular filtrate then travels through the *proximal convoluted tubule*, the *loop of Henle* (descending and ascending limbs), the *distal convoluted tubule*, the *connecting tubule* and the *collecting duct* sequentially. In the early stages of the proximal tubule, the reabsorption of glucose, amino acids and bicarbonate, along with Na^+ takes place [2]. The main mechanisms of reabsorption are the active reabsorption and the process of diffusion and osmosis. Glucose, amino acids and proteins are almost entirely reabsorbed through active reabsorption, while ions are also actively reabsorbed but in controlled quantities through regulatory mechanisms like hormone secretion. Water is also reabsorbed from the tubular fluid through osmotic diffusion. Certain substances are either reabsorbed very poorly or not at all from the tubules. Such a substance is urea, an end-product of protein metabolism that should be repeatedly removed from the organism for the protein metabolism to continue. While water is osmotically reabsorbed, only around 50% of the urea is reabsorbed, as it is significantly less likely to pass through the tubular membrane. Other substances that are treated in the same way as urea are creatinine, phosphates, sulfates, nitrates, uric acid and phenols [12].

Certain substances are actively secreted into the tubular fluid from the blood through the tubular epithelium. Such substances are potassium ions, hydrogen ions, ammonia, toxic substances and some drugs, for instance penicillin. Finally, the fluid that remains, urine, passes through the collecting duct into the renal pelvis and then through the ureter to the bladder [12].

2.3 Proteinuria

2.3.1 Diagnostic value of proteinuria

Proteinuria is widely known to be an indicator of different types of disease or abnormalities of human bodily functions. When detected it can be associated with glomerular disease, such as diabetes and cardiovascular diseases, urinary tract infection or lesion and renal disease, such as chronic kidney disease (CKD) [13], [14]. In certain cases, it has been linked to viral infections such as COVID-19, Epstein-Barr virus, hepatitis B and C viruses, human immunodeficiency virus [14].

During pregnancy, proteinuria was considered a sufficient symptom in order to diagnose preeclampsia. However, recent findings have shown that measurement of protein excretion in pregnant women alone should not be used to predict maternal and neonatal outcomes, but it remains a reliable method in diagnosing with preeclampsia women with CKD, chronic hypertension, diabetes or renal disorders [15]. Among children, some of the various infections that can cause proteinuria are a type of streptococcus infection, infectious mononucleosis or the presence of toxins (mercury, lead, copper) [16]. In the elderly, the most common causes of proteinuria are type 2 diabetes mellitus, multiple myeloma and amyloidosis [17].

2.3.2 Urine composition and types of proteinuria

The major proteins which constitute the urine of a healthy individual are the Tamm-Horsfall protein (THP) (50%), albumin (20%) and immunoglobulins (5%) [18]. Proteinuria is defined as urinary total protein excretion of more than 150 mg per day (about 10 to 20 mg per dL) [8], [19]. Proteinuria can be either transient, in which case the condition is classified as benign, or persistent, which is a sign of pathological conditions such as the ones mentioned in section 2.3.1. The most common type of persistent proteinuria is glomural, where albumin is the primary urinary protein. In tubular proteinuria low molecular-weight proteins are in excess in the urine and less than 2 g of protein is excreted during 24 hours. In overflow proteinuria, increased production of low molecular-weight proteins, such as Bence-Jones proteins or myoglobins, disrupts the nephron's ability to reabsorb filtered proteins which results in protein spilling in the urine [8].

As previously stated, a normal value of total protein excretion is less than 150 mg per day [13]. However, different guidelines suggest different cut-off values for abnormal

proteinuria, ranging between 150 and 300 mg/24h [20]. Low-grade proteinuria is 1 to 2 g/24h, a mild protein excretion, and requires follow-up testing to determine whether orthostatic proteinuria is present (a type of proteinuria that usually disappears with time). Nephrotic-range proteinuria is 3.5 g/24h, which in most cases signifies a severe glomerular disease [13]. Given that the average urine excretion is one liter per day, 150 mg/L can be considered a cut-off value for a non-proteinuric sample.

2.3.3 Measurement of proteinuria

The first step towards measuring proteinuria is defining which urine sample is most suitable for this purpose. The 24-hour collection is commonly considered the standard procedure [20]. Unfortunately, this is a time-consuming process, often inconvenient for the patient. On these premises, the 24h procedure is usually substituted by the measurement of the protein-to-creatinine ratio (UPCR) in a single spot urine sample [21]. The UPCR method's correlation to 24h urine protein (UP) measurement is supported by many researches [22], [23], [24], [25]. Its reliability stems from the fact that the urine protein excretion rate is proportional to the urine creatinine excretion rate, thus their ratio remains relatively constant through the day [26]. Because of variances of the urine creatinine (UCr) excretion among individuals, it is suggested to use the estimated UCr value obtained by the Cockroft-Gault equation [27] : 24h UCr (g) = [140 - age (years)] x weight (kg) / 5000 (x 0.85 in females) [26], [28].

A common means to detect proteinuria at the medical facility is the 'dipstick', a urine reagent strip device with a pad that changes color according to the level of proteinuria detected in the sample. Each manufacturer links the color change that does or does not occur with a certain range of urine protein found in the solution, while proteinuria in this case is sometimes defined as the color change attributed to a range of 300 mg/L of total protein or more [20]. Despite its convenience, the dipstick's sensitivity is generally low for the detection of proteinuria. The test is prone to false negative results when the urine is dilute or false positive results when the urine is concentrated. More importantly, due to the pH-sensitive dye in the dipstick, negatively charged proteins like albumin are most likely to be picked up by the device, whereas positively charged proteins like immunoglobulin light chains may not be detected even if they are in high concentrations [9]. As a result, weariness has spread among experts regarding the dipstick's ability to rule in or rule out significant proteinuria and most clinical guidelines either advise against testing with reagent strips or suggest follow-up laboratory testing,

given that this method often fails to detect some forms of kidney disease in the early stages and should be unsuitable for monitoring patients with known proteinuria [20].

Precipitation techniques are also employed for qualitative or semi-quantitative measurements in urine samples [14]. In order to do so, acid is added to the solution and the turbidity caused by protein precipitation is measured with a photometer or nephelometer. Different types of acid can be used to detect either albumin or globulins [9]. However, this method cannot detect protein levels below 50 mg/L [14]. Electrophoresis can be used to detect immunoglobulins and has been able to detect Bence-Jones proteins with sensitivity better than 1 mg/L [29]. This method requires a reference substance to measure protein on its own, otherwise it can function as a seperation technique to prepare the sample for mass spectroscopy, fluorescent spectroscopy and etc [14].

Mass spectroscopy is a research-lab method for biomarker discovery based on the measurement of the mass-to-charge ratio of ions. Combined with other techniques, it has identified more than 3000 individual proteins in the human urine. However, it does not provide a quantitative analysis in itself and to implement such features has proven to be a challenging task. Immunoassay, on the other hand, is a method used in both research and clinical laboratories [14]. This technique is dependent on a specific binding reaction between an antigen (analyte) and an antibody and can be performed in urine and other complex biological matrices without sample pretreatment [30]. In general, the immunoassay method has achieved lower limit of detection well below 1 mg/L in many cases [14].

Before further analysing the protein measurement techniques which will be used in the course of this study, the dilemma proteinuria or albuminuria measurement should be addressed. Total urine protein measurement suggests a challenging task because of the wide variety of protein concentrations and compositions found in different samples. Moreover, the amount of non-protein interfering substances is high comparing to the amount of proteins, affecting the efficiency of total protein measurement methods. Albumin, however, is the urine protein most likely to be picked up by the dipstick, turbidimetric and dye-binding methods. Measurement of urine albumin usually takes place using a specific immunoassay approach and as a result albumin assays achieve better analytical performance than total protein assays at lower concentrations (usually below 500 mg/L). Urinary albumin is widely used to detect early stages of diabetic nephropathy and glomerular pathology associated with hypertension, systemic sclerosis

and other systemic diseases. Microalbuminuria is a condition where total urinary loss is within non-proteinuric levels (below 150 mg/24h) and total albumin loss ranges between 30 and 300 mg/24h. It is a common symptom among non-diabetic individuals, correlated to cardiovascular and non-cardiovascular mortality [20].

It has been a common conclusion that total urine protein measurement cannot be achieved with a sole measurement technique. Electrophoresis can measure low-molecular weight proteins and enzymes, thus identifying tubular disease or Bence-Jones proteinuria, whereas total protein assays will not provide reliable results for such diagnosis. Urine albumin can be measured with specific immunochemical assays and, combined with selected protein and enzyme markers, could substitute total protein measurement [31].

2.4 Measurement techniques in focus

2.4.1 Spectrophotometric determination

In this section emphasis is given on the ultraviolet light absorption methods that are widely used for protein quantification. The basic principle behind absorption spectroscopy is that a photon is absorbed by an electron when the photon has an energy level matching the energy difference between two of the quantized energy states of the analyte [32]. The UV absorption spectrum of a protein can be considered as the sum of the spectra of the amino acids from which it is composed [33].

A widely used method to determine protein concentration is measuring the absorbance at 280 nm (A_{280}). It suggests one of the oldest methods and it is still widespread because of its simplicity and the fact that incubation with additional chromophores is not necessary. This technique requires that the proteins under study have aromatic amino acids, mostly tryptophan and tyrosine, that can absorb UV light [34]. At 280 nm, one absorbancy unit for a pathlength of 1 cm indicates protein concentrations of 0.5 to 2 mg/ml for a great number of globular proteins [33]. The equation:

$$concentration(mg/ml) = \frac{A_{280}}{a_{280} \cdot b}$$

can be used as a formula to calculate the unknown sample concentration, with a_{280} being the published absorptivity (ml/mg cm) at 280 nm and b being the pathlength (cm) depending on the instrument [34]. When the absorptivity a_{280} is unknown, the

concentration can be calculated by using the Beer-Lambert law:

$$A_{280} = \epsilon_{280} lC$$

where ϵ_{280} is the molar absorption coefficient (1/M cm), l is the pathlength (cm) and C is the protein concentration (M), provided that the value of ϵ_{280} is known. Although it is safer to measure ϵ_{280}, methods have been developed in order to predict it. For instance, the value of ϵ_{280} for a folded protein in water can be predicted by the equation:

$$\epsilon_{280}(1/Mcm) = (tryptophan) \cdot 5500 + (tyrosine) \cdot 1490 + (cystine) \cdot 125$$

Calculated values are reliable for proteins containing tryptophan, however not as much for proteins which do not [35]. Tryptophan and tyrosine in proteins can be measured with the Edelhoch method [35], [36]. Factors like the pH of the buffer, its polarity and ionic strength, might impact the absorbance spectrum of a protein as they can alter its structure, whereas buffers and their components in some cases interact with certain amino acids, further modifying the spectrum [32].

Protein quantification can also take place at 205 nm by measuring the peptide bond absorption. The A_{205} method can detect more types of protein than the A_{280} method and is able to quantify even lower concentration of protein because of the large number of peptide bonds in every protein. However, some buffers and other components, especially those that contain double bonds between carbons or carbon and oxygen, absorb at 205 nm, thus rendering the technique unreliable at times [34], [32]. In general, to prevent unwanted interference, the buffers should not be stored in plastic containers since they might release plasticizers that absorb at UV wavelengths, while the detergents selected need to have their carbon double bonds converted to single bonds so as not to absorb UV light [32].

2.4.2 Dye-Binding techniques

A very simple dye-binding technique was described by Bradford [11] who used as a reagent the Coomassie Brilliant Blue G-250, a red dye that converts to blue upon binding of the dye to the protein. It suggests a rapid, sensitive and easily implemented method to measure protein in a liquid solution. The basic mechanism of the Bradford assay is the binding of the dye at acidic pH to certain protein residues (tryptophan, tyrosine, arginine etc.), which leads to a metachromatic shift from 465 to 595 nm when

the anionic form of the dye becomes stable [37]. Specifically, Compton and Jones [38] have noticed that the dye reagent absorbance spectrum shows maxima at 470 and 650 nm that can be modified with the addition of certain substances to the reagent. For instance, addition of sodium hydroxide decreases absorbance at 465 nm, increases absorbance at 650 nm and then replaces the 650 nm peak with a new one at 595 nm, while concentrated sulfuric acid eliminates the peak at 650 nm, giving a single maximum at 465 nm. It is concluded that blue color is incident of the anionic form of the dye, with an absorbance maximum at 595 nm, green color suggests the neutral form (singly protonated species) of the dye with maximum at 650 nm and red color is associated with the cationic form (doubly protonated species), showcasing a spectrum peak at 465 nm [38].

A standard procedure for the protein measurement with Bradford assays involves the estimation of the increase of absorbance (A_{595}) against a reagent blank (dye reagent plus buffer) at a certain time window after the addition of the dye to the sample [39]. In detail, upon addition of protein to the dye reagent, a decrease in the 465 nm spectrum peak occurs, combined with a new broader height at 595 nm and a shift at 650 nm. When the dye blank spectrum is subtracted, the prevalent peak remains at 595 nm [38]. This subtraction can take place by setting to zero the spectrophotometer reading of the protein sample versus the blank at 465 nm before reading directly the absorbance at 595 nm and specifically aims to enhance the procedure's sensitivity for very low concentrations (below 10 μg/ml) [40].

Given that the received signal stems largely from arginine-related reactions, there is wide inner protein-to-protein variation in the Bradford assay. As a result, it is highly suggested that the sample and the standard are matched when this method is employed [37]. Alterations of the standard method that can increase sensitivity, comparability to other assays and reproducibility could be the addition of sodium dodecylsulphate (SDS) to the sample (aids detection of free light chain protein and IgG immunoglobulin) [41], [42], the use of a protein calibrator, such as a urinary protein lyophilizate (UPL) calibrator [43] and increasing the dye concentration while decreasing the phosphoric acid concentration [44].The results of the Bradford assay testing are generally not affected by the presence of nicotine, caffeine, several nitrogenous compounds, polycytidylic acid, adenine, purine [38].

Another issue that needs to be dealt with when using the Bradford assay is the non-linear relationship between the absorbance and the protein concentration that occurs in

certain concentration ranges. Bradford attributes this nonlinearity to the reagent itself, since there is an overlap in the spectrum of the different color forms of the dye and the background value for the reagent consistently diminishes as additional dye becomes bound to the protein [11]. Further research has shown that correcting the response curve using this claim does not fully linearize it, indicating an additional cause for the nonlinearization, especially for larger protein concentrations. A theoretical linear relationship between the ratio A_{590}/A_{450} and total protein concentration $[P]_t$ is given by the equation:

$$\frac{A_{590}}{A_{450}} = \frac{\epsilon_{590}^{DP} \cdot K \cdot f_b \cdot n}{\epsilon_{450}^{D}} \cdot [P]_t + \frac{\epsilon_{590}^{D}}{\epsilon_{450}^{D}}$$

where D stands for dye, P for protein, DP for the dye-protein complex and K is the equilibrium constant, with the equilibrium reaction being:

$$D + P \overset{K}{\leftrightarrow} DP$$

and the equilibrium constant being defined as:

$$K = \frac{[DP]}{[nP] \cdot [D] \cdot f_b}$$

where the [] symbol is used to indicate concentration, n is the number of binding sites on a protein molecule and f_b is the fraction of the reactive blue dye form (equals 0.053 under assay conditions). The values of ϵ_{450}^{D}, ϵ_{590}^{D} are a weighted average of the molar absorption coefficients of the three dye forms and are not affected by the addition of protein. This theoretical equation predicts that the linear curve will intercept the Y axis at the value $\epsilon_{590}^{D}/\epsilon_{450}^{D}$, which has been verified by experimental data [45].

The Bradford assay presents high sensitivity and wide detectable limits, making it one of the most suitable assays to detect protein in samples with very low protein concentrations such as the urine [46]. In order to measure urine protein using the Bradford assay certain issues must be taken into consideration. Given that urine is a complex biological environment, potentially containing other substances that affect the absorbance spectrum, the methods and equations described earlier might need to be modified. Afkarian et al. [47] have implemented the isobaric tags for relative and absolute quantitation (iTRAQ) method in protein identification and quantification for urine biomarker discovery. The urine processing protocol developed suggested that (1) the minimum required quantity of starting protein is 20 μg per sample, (2) protease inhibitors are not necessary for samples which are handled properly, meaning, for

instance, that they are stored at -80°C without repeated freeze-thaws and (3) an extra concentration step is usually required because of the variable and typically low protein concentrations, so organic solvent precipitation with methanol yields the most reliable and reproducible results when using the Bradford assay [47].

Other dyes that have been used to measure protein concentration are the bromphenol blue (BPB), a dye specifically bound by albumin [48], the pyrogallol red-molybdate (PRM) [43], the eosin dyes (eosin B and Y) [49], the bromocresol green (BCG) [50] and the bromocresol purple (BCP) [51] and the Pierce 660 nm protein assay based on the binding of a dye-metal complex to protein [52].

2.4.3 Colorimetry on smartphone

Colorimetry is the scientific and technological discipline employed to quantify and specify the human color perception. Digital Image Colorimetry (DIC) is an inexpensive, quick and robust tool that analyzes the concentration of the target analyte by visual examination of the color changes of the captured digital image [53]. One of this method's purposes is to eliminate the subjective error of the naked-eye observation that unavoidably interferes with the measurement process and to establish a uniform classification different color systems, or color spaces, are employed. A color space represents a collection of colors detectable by a human or a device. When referring to a screen, the color space is a sum of all the colors it is able to display. There are several color spaces such as RGB, CMYK, HSV, XYZ, L*a*b* and the Gray model, with RGB, HSV and Gray the most widely used [54]. A brief overview of the topic is presented here, whereas supplementary information on the field of digital image processing and the concept of colorimetry is provided in Appendix B.

In computer-based display systems the color space is often visualized by a unit cube and each one of the axes in the 3D system is assigned to a color (**R**ed, **G**reen and **B**lue). The arithmetic value of each axis suggests the intensity of each component and ranges from zero contribution, to full or maximum contribution to the color. Different devices assign different arithmetic values to components, with the most common being 0-255. Alternatively, the HSV color space is represented by a cylindrical coordination system and stands for **H**ue, the angular dimension that yields the color type (red, blue or yellow for instance) ranging from 0° to 360°, **S**aturation, the radius of the cylinder receiving values from 0 to 100% that shows the intensity of the color and **V**alue, or **B**rightness (HSB), the longitudinal axis that depicts the brightness of the color, ranging from 0 to

100% [54].

The perceived color of images can be influenced by multiple factors concerning the lighting conditions, including the uniformity and the intensity of the light source, the roughness of the object surface and the homogeneity of the colorization. Due to these types of interference, the ideal situation would be for the analyte to be photographed under exactly the same lighting conditions [54]. In researches that have tried to implement colorimetry in a camera-obtained image, a common method has been to use a fixed lighting environment, for instance a closed device and three symetrically placed LED beads as illumination sources [55] , or two LED beads at 45 degrees angle to the sample [56], [54].

On the other hand, software solutions have also been developed to resolve illumination-related issues by developing algorithms for color correction. For example, Hong and Chang [57] developed an algorithm that corrects each of the RGB components of a specific image portion, called region-of-interest (ROI), using the following formula:

$$R_{corr} = \frac{256}{R_W - R_B}(R_{meas} - R_B)$$

where R_{corr} is the corrected value, R_W the white background value, R_B the black background value and R_{meas} the measured value of the sensor. In this case, the algorithm measures the black and white background values it can detect in the image, essentially establishing a different color scale for every different environment lighting. Jia et al. [58] utilized the aforementioned formula to create a calibration technique that accounts for variations that occur from light conditions, imaging positions and cellphone brands. This technique proved successful in the context of detecting glucose in urine using a colorimetric sensor array.

Mathaweesansurn et al. [59] have developed an application that measures urinary albumin based on the reaction between albumin and tetrabromophenolphthalein ethyl ester (TBPE), which gives the analyte a distinct green color. The acquired photo consisted of a standard colorimetric strip containing the six reference colors (multiple shades of green each attributed to a different albumin concentration, ranging from 1 to 50 mg/L) and the two sample holders, one being the control holder (containing a quantity of the urine sample in deionized-distilled water) and the other being the test holder (containing the same quantity of urine sample incubated with TBPE in acetate buffer). The calibration in this case takes place by calculating and comparing the H parameter in the HSV color space and manages to eliminate illumination-oriented

variations without any additional lighting control apparatus.

Another way to measure protein concentration through DIC has been proposed by Moreira [60], utilizing the RGB color space. The blue to green ratio of a sample enriched with Bradford assay was used to accomplish improved linearity and greater range by applying the corresponding equation that was mentioned in section 1.4.2. Camargo et al. [61] have also used the Bradford assay for the same purposes, only this time employing the HSV color space to provide the analytical signal (S_a) as such:

$$S_a = log\frac{H_s}{H_b}$$

where H_s is the value of the hue component of the sample and H_b of the blank accordingly and were acquired by directly converting the RGB components.

Chapter 3

Experiment Preparation

3.1 Materials and Equipment

The Bradford Reagent (5x) was purchased from SERVA Electrophoresis GmbH. The protein used was Bovine Serum Albumin (BSA) with 98% purity, purchased from Sigma Aldrich. Deionized water was also bought for use in sample preparation. All the urine samples used were taken from the same healthy subject. The urine sample was refrigerated and reused when the measurements required it, following the guidelines about urine sample treatment. In the sample preparation the instruments used were a precision balance for the weighing of BSA powder, pipettes and a vortex mixer, while for the measurement of color, the equipment used was a spectrophotometer, plastic wellplates, glass-microfiber paper and the test platform printed with a 3D-printer, as well as smartphones for the image acquisition.

3.2 Experiment procedures

3.2.1 Initial Experiments

The first experiments were performed in order to gain familiarity with certain techniques and further develop the research questions. In these experiments different quantities of BSA solutions were prepared and mixed with the Bradford solution to determine the most effective way to extract the calibration curve. In three consecutive sessions, three batches of solutions with different ranges of protein concentration were measured with the spectrophotometer to determine their absorbance at 595 nm. The Bradford solution was prepared by adding 5 ml of the purchased Bradford reagent (5x) to 20 ml of distilled water, thus producing a 1x concentrated Bradford solution. The different protein solutions occurred by taking various quantities from a BSA stock solu-

tion of 0.8 mg/ml and adding distilled water to reach a specific volume, same for all the solutions. The measurements were performed in duplicates. The three concentration ranges selected were: 1-5 μg/ml, 5-25 μg/ml and 100-1000 μg/ml.

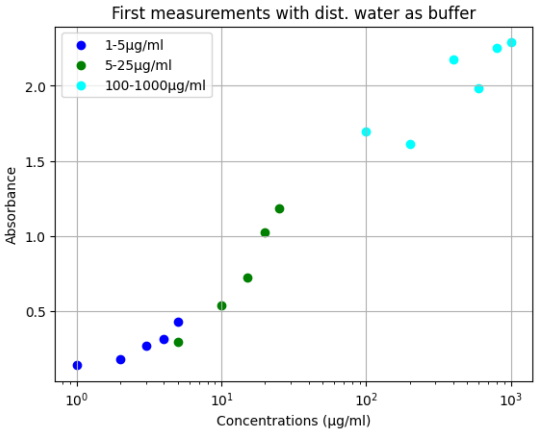


Figure 3.1. *Results from the initial experiment*

These measurements' results are depicted in Figure 3.1, where the x axis is in logarithmic scale. It can be observed that the relationship between absorbance and protein concentration has good linearity for each of the 1-5 and 5-25 μg/ml ranges, whereas for higher concentrations this method is rather insensitive and yields poor results. This behavior is also demonstrated in Figure 3.2, where a first-degree polynomial line is drawn for each range after applying linear regression.

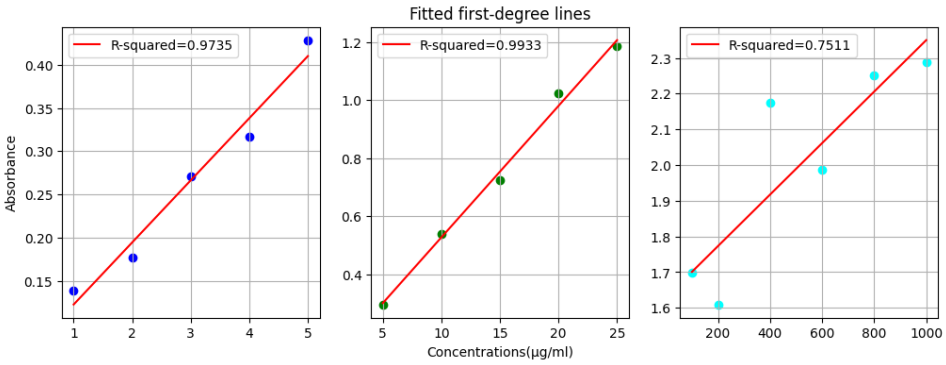


Figure 3.2. *Linearity of the method in different concentration ranges*

The proposed sensing method is developed for urine protein quantification. Therefore, it is suggested that while creating the calibration curve, the buffer with which the protein standards are diluted should match the buffer of the protein samples (in this case, urine). As a result, the following experiment was conducted, with different volumes of BSA solution (2,4,6,8 and 10 μl) diluted in the same volume of urine (2 ml). The BSA solution was prepared by weighing 5 mg of BSA and diluting it in 1 ml of dist. H_2O.

This procedure was used to prepare samples with respective protein concentrations of 5, 10, 15, 20 and 25 μg/ml approximately. Then, the following assumption was made: if the linearity of the method is improved for lower sample concentrations, as indicated by the previous experiment,it is possible that a smaller quantity of BSA could be measured with the spectrophotometer. Afterwards the measured absorbance would have to be multiplied by a certain factor, assuming again linearity in the response curve.

Therefore, I took 400 μl from each of the previous samples (urine spiked with BSA) and added each to 1.6 ml of the Bradford 1x solution. The spectrophotometer was zeroed at 595 nm against a blank sample consisting of 200 μl urine and 800 μl Bradford solution. The measurements were performed in duplicates. The amount of protein measured in this case is equivalent to one fifth of the original sample's content. Hence, in order to calculate the corresponding absorbance to the original sample concentrations, I multiplied the received absorbance values by five.

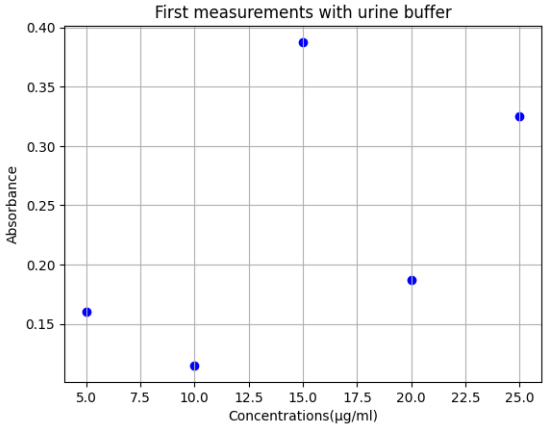


Figure 3.3. *Results of the described multiplication method*

As shown in Figure 3.3, no correlation was noticed between absorbance and protein concentration in this instance. In order to ensure that this was not due to some sort of mishandling, the same experiment was repeated twice, with some modifications in the concentrations of the original BSA solution, which in theory should leave the actual results unaffected.

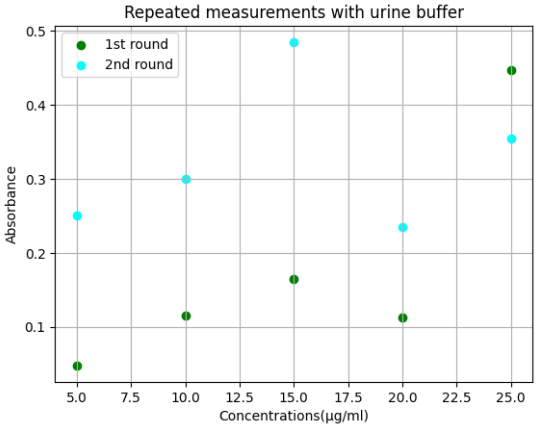


Figure 3.4. *Results of the replicated measurements using the multiplication method*

As is also demonstrated in Figure 3.4, the proposed multiplication method is obviously ineffective. Its proposition in the first place came under the assumption that the urine used as buffer would already contain a sufficient amount of proteins and other big molecules. On one hand, one would expect these to create a substantial shift in the samples' color when mixed with the Bradford solution. On the other hand, one would believe that the size of the molecules might also cause an overall significant drop in the absorbance measured with the spectrophotometer.

The previous assumptions were eventually not validated by these results. In fact, the addition of the undoped urine sample in the Bradford solution resulted in only a minor shift in color, as visually observed, implicating that the urine originally used did not contain a sufficient amount of protein to "push" the measurements out of the linear range. Additionally, by zeroing the spectrophotometer at 595 nm every time with a blank sample consisting of only the Bradford solution and the urine sample used throughout the rest of the measurements (in the same ratio as in the rest of the samples) without spiking it with additional BSA, the subtraction of the initial protein quantity from the measurements can be achieved.

In order to validate this hypothesis, I performed a series of measurements using urine as buffer. The different ranges of concentration values measured were: 50-100, 12.5-62.5, 5-10 μg/ml. Initially, 50 mg of BSA powder was measured and dissolved in 1 ml of urine. This solution was then dilluted with another 9 ml of urine, thus leading to a 5 mg/ml solution (BSA+urine). Throughout the process, 1x Bradford solution was used, prepared in the same way as previously described. In 2 ml of the 1x Bradford solution 20, 25, 30, 35 and 40 μl of BSA+urine was added to create concentrations of approximately 50, 62.5, 75, 87.5 and 100 μg/ml respectively. For the blank sample I

used 2 ml of Bradford and 40 μl of undoped urine. In the second batch, I took 5, 10, 15, 20 and 25 μl from the same urine+BSA solution, adding each one to 2 ml Bradford, receiving solutions of 12.5, 25, 37.5, 50, 62.5 μg/ml approximately. For the blank sample I added 25 μl of urine to 2 ml of Bradford. In the final batch, I further diluted 1 ml of the original BSA+urine in 9 ml urine, which results to a 0.5 mg/ml solution. I then add 20, 25, 30, 35 and 40 μl in 2 ml of Bradford to achieve roughly concentrations of 5, 6.25, 7.5, 8.75, 10 μg/ml. For the blank I used 2 ml Bradford with 40 μl of urine. The measurements were all performed in duplicates. The results are depicted in Figures 3.5 and 3.6.

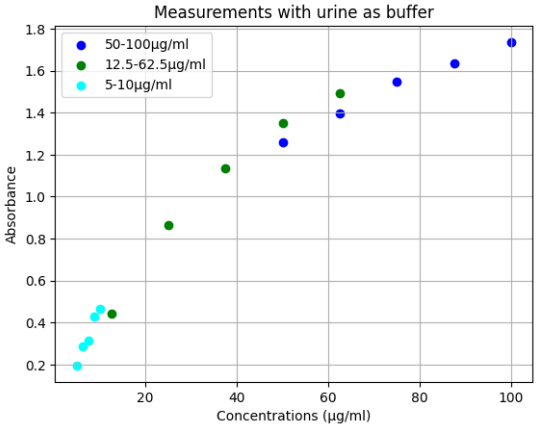


Figure 3.5. *Results acquired with the established method in urine buffer*

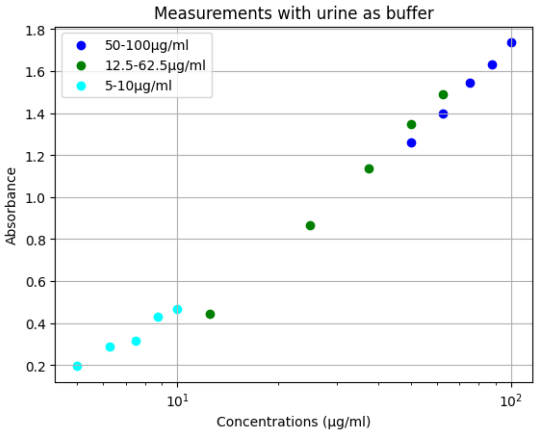


Figure 3.6. *Results acquired with the established method in urine buffer (logarithmic scale in the x-axis)*

It is noticed that the linearity increases when the concentrations are plotted in the logarithmic scale. Nevertheless, to test the linearity in the different concentration ranges, linear regression was used again, as shown in Figure 3.7.

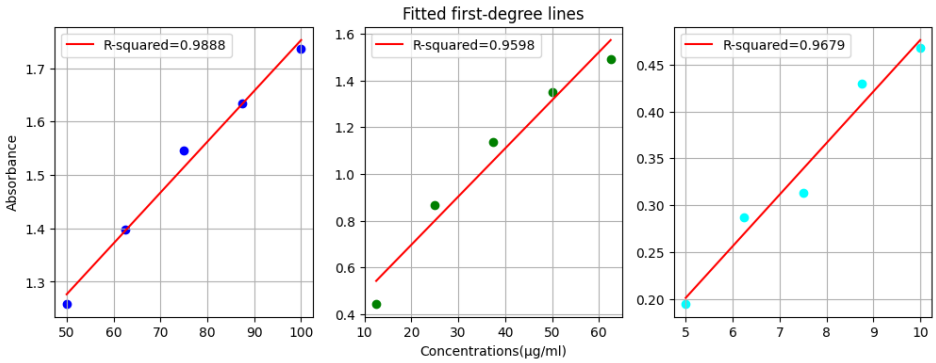


Figure 3.7. *Linearity in different concentration ranges*

It can now be concluded that it is possible to create a sufficiently linear calibration curve for a range of approximately 5-100 μg/ml protein concentrations. The questions that now occur are whether **(a)** the linear range can be expanded, **(b)** the linearity can be improved by experimenting with different Bradford to sample ratios, and **(c)** a relationship between a dist. water calibration curve and a urine calibration curve can be found.

3.2.2 Creating the calibration curve

The Bradford assay that was purchased is able to measure protein concentrations from 1 up to 1000 μg/ml according to the instruction manual. The linearity of the assay, however, is not guaranteed for such broad range, as seen in the measurements performed with dist. water as buffer (Figure 3.2). In order to test the linearity specifically for measurements in urine, the previous procedure was repeated.

Firstly, I measured 50 mg of BSA and dissolved it in 1 ml of the urine sample. From this 50 mg/ml BSA+urine solution I pipeted the quantities required to prepare the macro-concentrations. The 1x Bradford solution I used is the same throughout this experiment (8 ml of 5x Bradford reagent mixed with 32 ml of dist. water). In 2 ml of 1x Bradford I pipeted 20, 24, 28, 32, 36 and 40 μl BSA+urine to approach concentrations of nearly 500, 600, 700, 800, 900 and 1000 μg/ml respectively. To proceed to medium concentrations, I need to further dilute the remaining BSA+urine solution until a concentration of 5 mg/ml is reached. For this purpose, 7.38 ml of urine was added to the remaining 0.82 ml of BSA+urine. From this solution I pipeted 20, 40, 60, 80 and 100 μl into 2 ml of 1x Bradford to achieve sample concentrations of approximately 50, 100, 150, 200 and 250 μg/ml respectively. Finally, to accomplish even lower sample concentrations, I formed a 0.5 mg/ml BSA+urine solution by diluting

1 ml of the previous 5 mg/ml BSA+urine solution with 9 ml of urine. Next I added into 2 ml of 1x Bradford, 4, 20, 40, 60, 80 and 100 μl of BSA+urine to approximately reach concentrations of 1, 5, 10, 15, 20 and 25 μg/ml respectively. For all the measurements I used the same blank sample, consisting of 2 ml Bradford and 100 μl of urine. All the samples were mixed in the vortex and incubated for 5 minutes in room temperature. The measurements in the spectrophotometer were performed in duplicates.

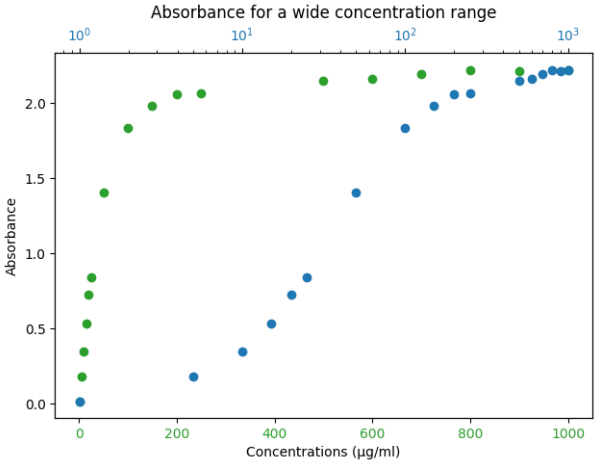


Figure 3.8. *Measuring a wide range of protein concentrations (green-linear and blue-logarithmic scale in the x axis)*

Simply by observing the results in Figure 3.8, certain estimations can be made about the linearity of the method. It can be safely deduced that concentrations above 400 μg/ml cannot be measured with sufficient accuracy with the technique proposed so far. To further test the linearity I followed the series of calculations depicted below in Figure 3.9.

The method has significant linearity for concentrations leading up to 50 μg/ml. The linearity is sufficient up to 100 μg/ml but is then declining. A factor that can potentially lead to an increase in linearity is the concentration of Bradford assay with which the sample protein is measured.

In this series of experiments the samples were tested with multiple Bradford solution concentrations. In the first experiment, the measurements were performed with 10% and 40% Bradford assay concentrations. In both cases, in 2 ml of 10% or 40% Bradford solution, 20, 40, 80, 120, 160 and 200 μl of 5 mg/ml BSA+urine were added to reach sample concentrations of approximately 50, 100, 200, 300, 400 and 500 μg/ml respectively. In the first case, the 10% Bradford solution was prepared by mixing 1.6 ml of the 5x Bradford assay with 14.4 ml of dist. water, resulting to a solution with a light blue color. In the second case, the 40% Bradford solution was prepared by adding

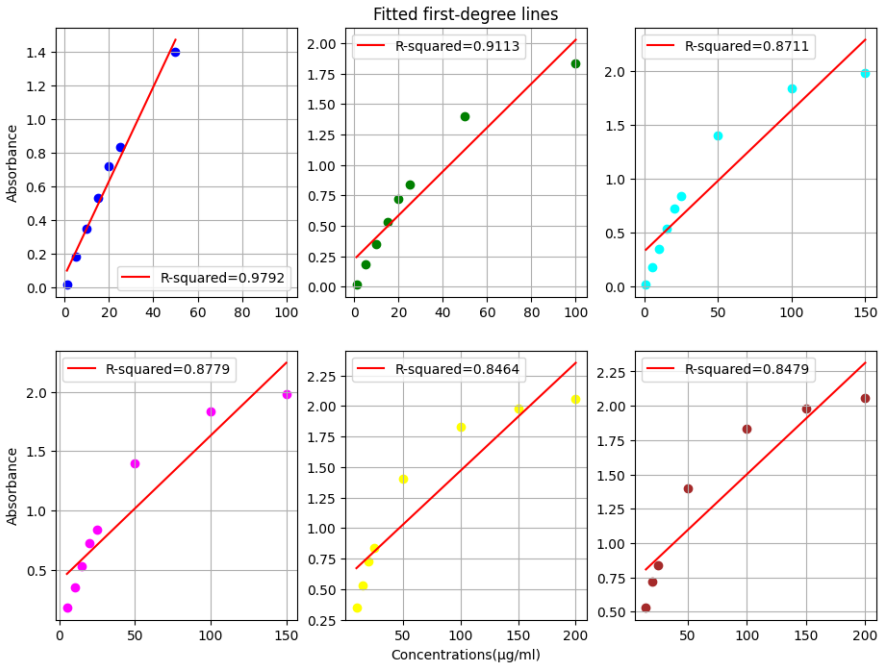


Figure 3.9. *Testing the linearity in different concentration ranges*

6.4 ml 5x Bradford assay to 9.6 ml of dist. water, which gave the final solution a deep red-brown color. The blank sample for each set of measurements consisted of 2 ml 10% or 40% Bradford solution and 200 μl urine. The measurements were all performed in duplicates.

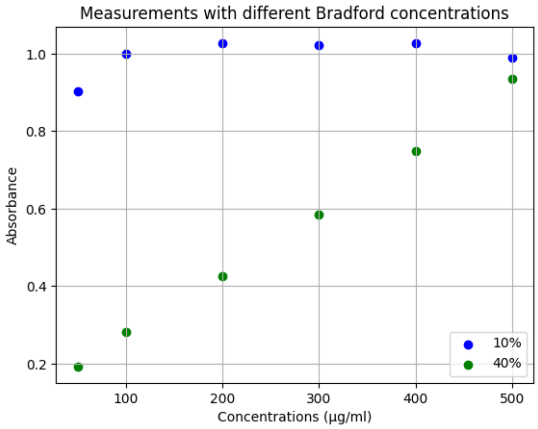


Figure 3.10. *Examining if adjusting the Bradford concentration might affect the linearity of the method.*

With the increase of Bradford concentration, the linearity of the method is improved, as shown in Figure 3.10, but the overall absorbance values are significantly lower than the values previously noticed. This is specifically concerning given that the method developed is destined for phone camera detection and so the accuracy levels provided by the spectrophotometer cannot probably be reached.

In the second experiment, the Bradford solutions prepared had concentrations of 5, 10, 20 and 30%. To achieve these concentrations and reach a 20 ml volume solution I dissolved: 1 ml 5x Bradford assay in 19 ml dist. water (5%), 2 ml in 18 ml (10%), 4 ml in 16 ml (20%) and 6 ml in 14 ml (30%). In order to also achieve more precise sample concentrations, a 200 μl total volume of BSA+urine and urine would have to be added to 1.8 ml of Bradford solution. This way, 20 μl of 5 mg/ml BSA+urine was added to 180 μl of urine (50 μg/ml), 40 μl to 160 μl (100 μg/ml), 80 μl to 120 μl (200 μg/ml), 120 μl to 80 μl (300 μg/ml), 160 μl to 40 μl (400 μg/ml) and finally just 200 μl of BSA+urine (500 μg/ml). The blank sample consisted of 200 μl urine and 1.8 ml Bradford. The measurements were all performed in duplicates.

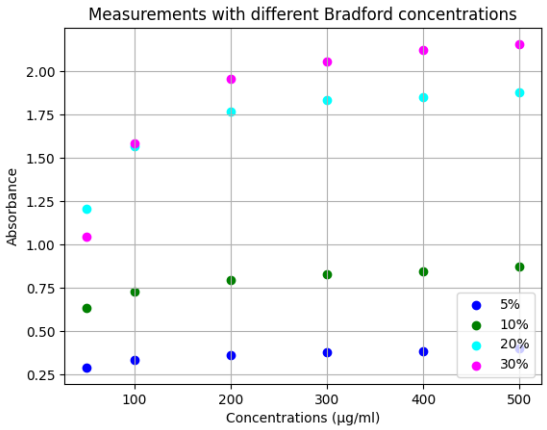


Figure 3.11. *Comparison between the response curves for different Bradford solutions.*

As shown in Figure 3.11, the sensitivity of the assay is increased for the Bradford concentration of 30%. For the purpose of the following experiments, a 30% Bradford solution is going to be used in accordance with these findings.

In the next step, a comparison between a dist. water and a urine calibration curve was performed. For both experiments the same Bradford 30% solution was prepared, by mixing 19.2 ml of Bradford 5x assay with 44.8 ml dist. water. In both curves the same procedure, shown in Table 3.1, was followed except for the difference in the buffer solution (in one case dist. water, in the other urine). The blank samples were prepared by adding 200 μl of buffer to 1.8 ml of 30% Bradford.

The results of the aforementioned procedure can be shown in Figure 3.12. It is evident that there is no significant difference between the two curves, even though the linearity in both curves is decreasing for higher concentrations.

To perform measurements based on a picture taken with a phone, the values I have so far will need adjustments based on the color space in which the analysis is going

Sample conc.(μg/ml)	25	50	75	100	125	150	175
Bradford sol.(ml)	1.8	1.8	1.8	1.8	1.8	1.8	1.8
BSA+buffer(μl)	10	20	30	40	50	60	70
Buffer(μl)	190	180	170	160	150	140	130
Sample conc.(μg/ml)	200	250	300	350	400	450	500
Bradford sol.(ml)	1.8	1.8	1.8	1.8	1.8	1.8	1.8
BSA+buffer(μl)	80	100	120	140	160	180	200
Buffer(μl)	120	100	80	60	40	20	0

Table 3.1. *Experiment procedure for the calibration curves.*

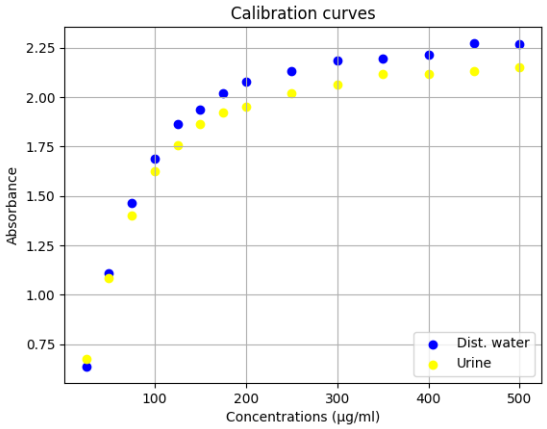


Figure 3.12. *Comparison between calibration curves in different buffers.*

to take place. This means that a common scale needs to be introduced to connect the absorbance values from the spectrophotometer with the RGB or HSV or L*a*b* values from the picture.

This kind of measuring system can be introduced by applying a normalization technique. The type of normalization that fits our kind of data is the linear normalization or "min-max" normalization. The scaling of the data is done according to the formula [62]:

$$X_{new} = \frac{X - min(X)}{max(X) - min(X)}$$

where X is the data value prior to the normalization, min(X) the minimum value of the dataset, max(X) the maximum value of the dataset and X_{new} the normalized data value. After the normalization, all the data falls in the range between 0 and 1, where 0 is the equivalent of the min(X) and 1 the equivalent of max(X), as can also be deduced by the former equation. When this normalization is applied to the former data, for each of the curves separately, the result is presented in Figure 3.13.

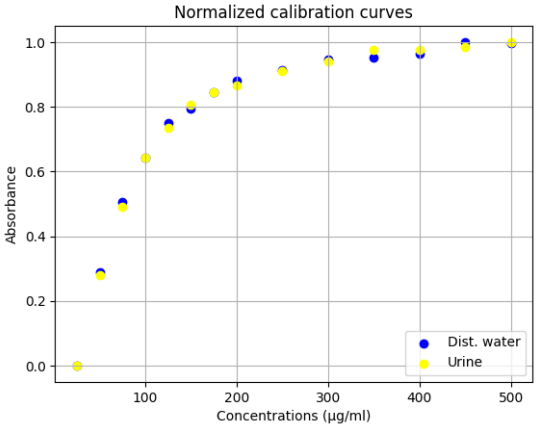


Figure 3.13. *The calibration curves after min-max normalization.*

It can be easily seen that the two curves post-normalization almost overlap. This fact might prove useful when also dealing with other buffers, potentially indicating that the measurements can be performed with reference only to a dist. water calibration curve.

3.2.3 First step of image color processing

In the following stage, the extent to which color can be inferred from images will be investigated. In particular, a 48-well plate was used to hold 8 different samples, prepared as shown in Table 3.2.

Sample conc.(μg/ml)	25	50	100	150	200	300	400	500
30% Bradford sol.(μl)	450	450	450	450	450	450	450	450
5 mg/ml BSA sol.(μl)	-	-	10	15	20	30	40	50
0.5 mg/ml BSA sol.(μl)	25	50	-	-	-	-	-	-
Buffer(μl)	25	-	40	35	30	20	10	-

Table 3.2. *Sample preparation for the 48-well plate.*

The Bradford solution used had a 30% concentration (as previously clarified), the first protein solution contained BSA in a 5 mg/ml concentration and the second protein solution was prepared by a tenfold dilution of the first solution, giving a 0.5 mg/ml concentration. In this step, the buffer was dist. water. It is also pointed out that the sample to Bradford ratio is the same with the one used in the previous measurements concerning the calibration curve.

Subsequently, a series of pictures were taken with a smartphone of the 48-well plate containing the samples. Out of these images, three were selected for further processing, taking into account the shading, the brightness and the reflections in the various images

(Figure A.1.1 - found in Appendix A). The colors in these images were analysed with the ImageJ software in three different color spaces, HSV, RGB and L*a*b*, in order to select the color space's channels that provide the most sensitive measurements, closest to the calibration curve.

The three images were uploaded sequentially to the ImageJ application where each RGB image (the default format for smartphone camera images) was split into the channels of the aforementioned color spaces. The colors in the channels after the split were represented in grayscale. In order to measure the intensity in each channel, an arbitrary rectangular region of interest was selected over every colored well in the plate. The mean values of all the pixels' intensities from each region were then plotted against the corresponding sample concentrations. The results of this analysis are presented in detail in Appendix A.

Then, for each color space, a characteristic channel was selected that seemed to follow a trend identical to the calibration curve. The selected channels from each color space were the blue channel from the RGB, the saturation channel from the HSB and the b* channel from the L*a*b* color space. In order to make direct comparisons between these values a common scale had to be established. Given that the lowest and highest concentrations were the same for both the calibration curve and this set of measurements, the min-max normalization technique could be used in this case as well. The results for this analysis are presented for all three images in Figure A.1.5. In Figure 3.14 the mean normalized values of the different pictures are presented to estimate which channel is the one closest to the calibration curve. The b* channel pf the L*a*b* color model seems to closely approach the results of the spectrophotometer.

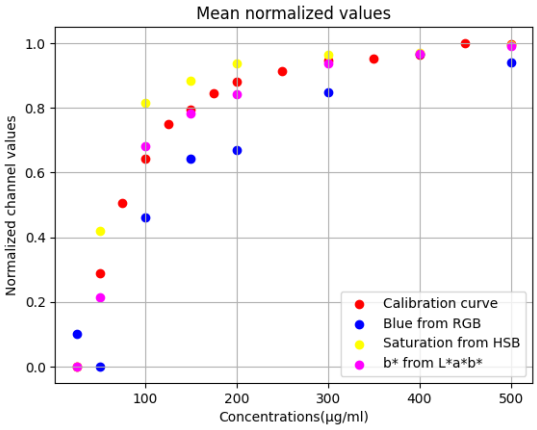


Figure 3.14. *Mean normalized values for all selected channels compared with the calibration curve.*

The problem with this method is mainly the error in the manner that pictures are taken. If the pictures are taken in normal lighting conditions (indoor lighting in a sunny day or with a light source on the ceiling) there are reflections on the surface of the sample. These reflections are also present when the picture is taken with the flash of the camera on, resulting to the inhomogeneity of the color area from which the measurement has to be extracted (Figures A.1.1b and A.1.1c). In order to minimize these reflections, a different, more complex setup of the camera and the plate can be examined. In Figure A.1.1a for instance, the picture of the bottom of the plate was taken while applying a white background on top of the plate to achieve better shading. While the reflection effect is significantly reduced, the picture becomes darker and in certain color models, the shades of blue become less distinguishable, such as the RGB color model (Figure A.1.2). This configuration also required a front-facing camera with good resolution, which might not be commonly available. A common problem in all instances is that the total area under measurement should become smaller for the circles to be more uniform and for the camera to be able to focus. Consequently, these possibilities were examined: **(a)** whether the measurement can be made on a different surface, a paper surface for instance and **(b)** whether the quantities can become smaller so that the total measurement area is reduced as well.

3.2.4 Measurements on paper

In order to perform measurements on a paper surface, a 30% Bradford solution was prepared by mixing 1.5 ml of 5x Bradford assay with 3.5 ml of dist. water. At first typical Whatman No.1 filter paper was used to test the response of the paper to different protein concentrations. However, the results were poor as the paper surface seemed to interfere with the assay. Such findings were also reported by Pokhrel et al. [63], who suggested that the interference is possibly caused by the cellulose in the paper and that glass-microfiber paper is used instead.

Firstly, certain preliminary, qualitative observations were made using the glass-microfiber paper. The paper disks were cut in small circles of 6 mm diameter with a paper punch. They were then placed inside a 96-well plate to be held steady (otherwise they were very lightweight and would move even by breathing close to them). Various volumes of 30% Bradford solution were added on the surface of the paper to examine how much liquid the paper can absorb, the quantities for improving the color display on the surface and the *coffee ring effect*.

The coffee ring effect is the result of the capillary flow that is caused during the evaporation of a liquid drop from a surface. The liquid evaporating from the edges is replenished by liquid from the interior, creating a flow that carries the dissolved material to the periphery. The flow is even more intense during the final stages of the evaporation process [64]. The coffee ring effect was particularly visible for liquid volumes up to 15 μl. For volumes above 20 μl excess liquid was present in the wells that could not be absorbed by the paper disks. In general, due to significant evaporation, it was concluded that for greater measurement stability and uniformity, larger volumes of reagents should be selected. It was also observed that the pipet should not be in direct contact with the paper when releasing the liquid to avoid the appearance of a more color-intense region at the point of contact.

In the following experiment different volumes and different Bradford solution concentrations were examined to choose the ratio that yields the best color response. This is a necessary step because the calibration curves and all the previous measurements have been performed with a much smaller sample-to-Bradford solution ratio. However, the same ratio cannot be maintained here because the sample volumes would become too small for pipeting to be reliable (there could be liquid excess on the surface of the pipet's tip that would not be accounted for). For this reason a minimum pipeting volume of 10 μl is established.

It should also be pointed out that the concentrations under measurement so far have been estimated by calculating the micrograms of a known BSA quantity in the total solution volume (the amount of BSA in buffer added to the Bradford solution). This formula can be now altered while approaching the final stages of the sensor development to assess the concentration of the BSA solution under measurement only. The following experiment illustrates this shift in formalism. The measurements take place by adding a volume of 10 or 30 μl of a BSA solution with known concentration to 30 μl of a Bradford solution with 30% or 50% 5x Bradford assay content. The different sample concentrations were prepared from a 0.5 mg/ml BSA solution as shown in Table 3.3. The glass-microfiber paper was cut with a paper punch into disks with a diameter of 6 mm and placed in a 96-well plate as also described before. The four rounds of measurements are shown in Figure 3.15.

By extracting the b* values of the L*a*b* model in ImageJ and applying the min-max normalization, the different color responses in Figure 3.16 were received:

Sample conc.(μg/ml)	40	100	200	300	400
0.5 mg/ml BSA sol.(μl)	50	50	100	150	400
dist. water(μl)	575	200	150	100	100

Table 3.3. *Preparing the protein samples used for every set of measurements.*

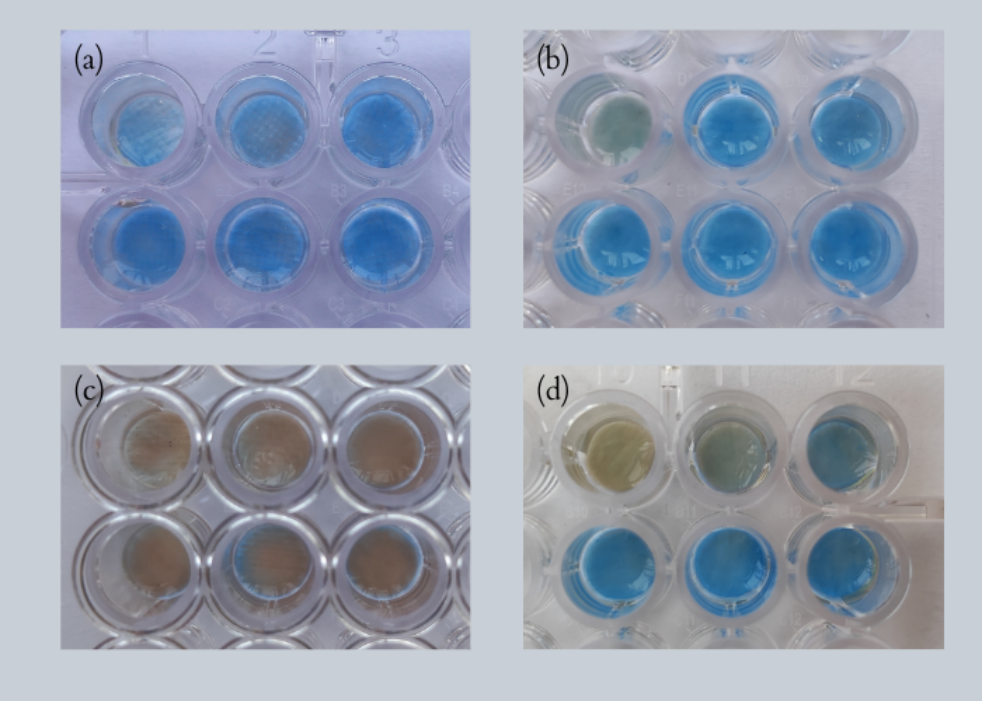


Figure 3.15. *The four sets of samples with each well containing in total:* ***(a)*** *10 µl of BSA solution and 30 µl of 30% Bradford solution,* ***(b)*** *30 µl of BSA solution and 30 µl of 30% Bradford solution,* ***(c)*** *10 µl of BSA solution and 30 µl of 50% Bradford solution,* ***(d)*** *30 µl of BSA solution and 30 µl of 50% Bradford solution. The well on the top left of each set contains the blank solution, the one on top right the 100 µg/ml sample, the one on the bottom left the 200 µg/ml and the bottom right the 400 µg/ml.*

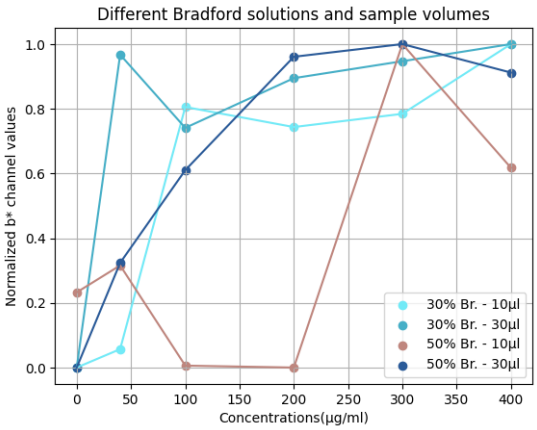


Figure 3.16. *The results of the color extraction for the different images of Figure 3.15. The parameter on the y-axis is the b* channel value normalized in a zero-to-one scale.*

Most of the lines in Figure 3.16 are not particularly close to the shape of the cali-

bration curve that has often reappeared through this section. However, the data points that have occurred for the set of samples with 50% Bradford solution and 30 μl follow a similar overall trend with the calibration curve, being almost linear for smaller concentrations. Because of this resemblance, these proportions will be kept for the rest of the experiments in the sample under measurement.

In this next step the 50% Bradford solution with the 30 μl Bradford-30 μl sample ratio is conserved, only this time a greater amount of concentrations will be estimated. Nine samples with different BSA concentrations were prepared as shown in Table 3.4:

Sample conc.(μg/ml)	0	20	50	100	150	200	250	300	350
0.5 mg/ml BSA sol.(μl)	-	40	100	200	300	400	-	-	-
5 mg/ml BSA sol. (μl)	-	-	-	-	-	-	50	60	70
dist. water(μl)	1000	960	900	800	700	600	950	940	930

Table 3.4. *Samples with different concentrations for estimation.*

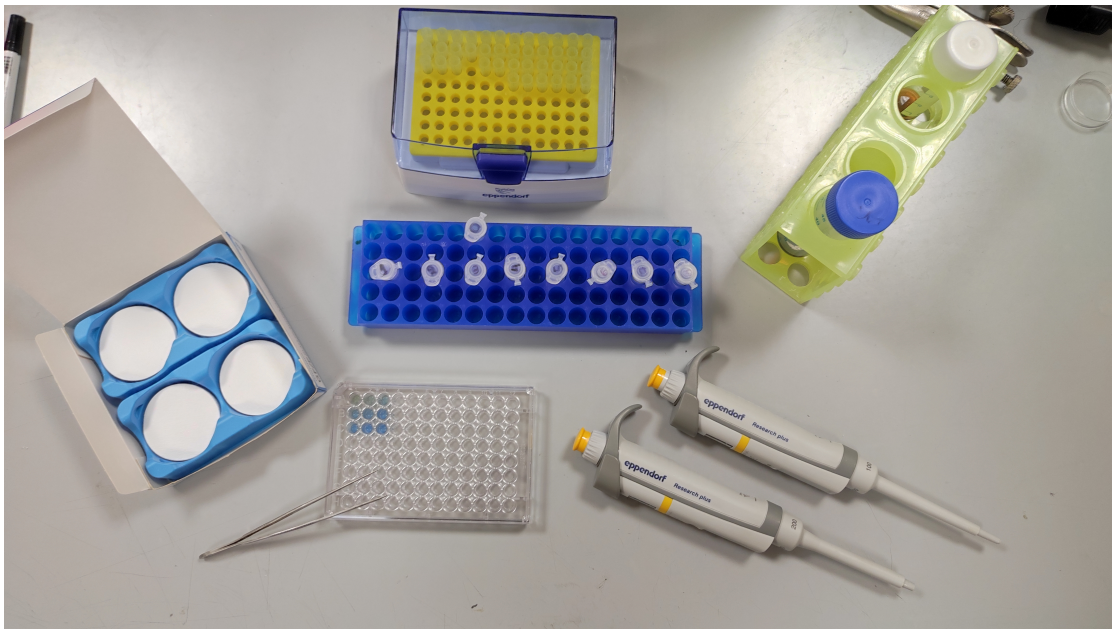


Figure 3.17. *The materials and setup used for the experiment of Table 3.4.*

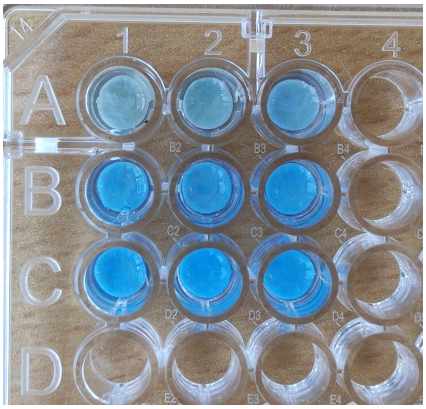


Figure 3.18. *The paper disks with the color response to the different protein concentrations. In the well denoted by A1 is the blank sample (zero protein content) with concentration increasing from left to right until the well C3 with a concentration of 350 μg/ml.*

The materials used for the experiment are shown in Figure 3.17. The glass-microfiber paper disks were cut into smaller circular disks with a paper punch in order to fit in the wells of the 96-well plate. In each well, 30 μl of the 50% Bradford solution was added along with 30 μl of a protein solution with a specific concentration, prepared as described in Table 3.4. Upon addition of the protein sample, the solution was slightly stirred with the tip of the pipet for color uniformity. Pictures of the colored samples were taken under different lighting conditions (shown in Figure 3.18 a picture taken close to the window with natural lighting). The colors in the different wells were estimated by extracting their b* value with ImageJ and their normalized values are plotted in Figure 3.19 against the respective sample concentrations.

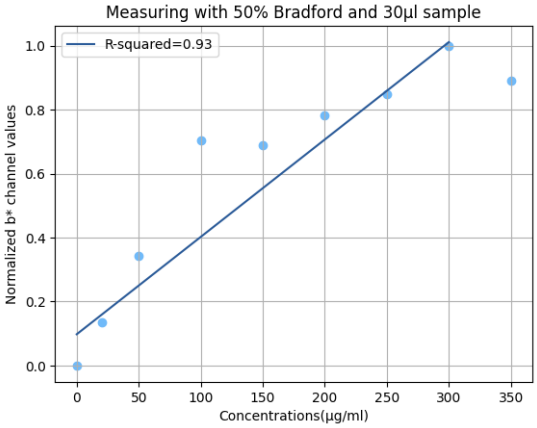


Figure 3.19. *The results of the color extraction for Figure 3.18. The y-axis results are the values of the normalized data. The line was fitted with all the data points minus the last one for 350 μg/ml.*

3.2.5 Designing the test platform

From the previous experiments and the whole process, two objectives emerged for more accurate color measurement: **(a)** depicting smaller regions of particular shape, minimizing local reflection and shading effects and **(b)** extracting color by determining its relative position in a standard scale. In this context, the test platform shown in Figure 3.20 was designed.

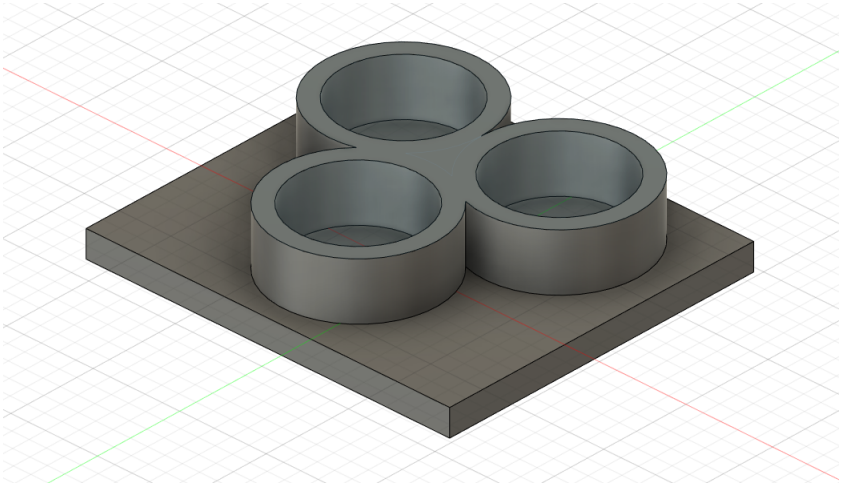


Figure 3.20. *The design of the test platform in the Fusion 360 environment.*

The test platform was designed with the Autodesk Fusion 360 program and printed in the lab's 3D printer. The test platform consists of three wells which were designed as hollow cylinders with an inner diameter of 7 mm (slightly larger than the diameter of the cut paper disks) and an outer diameter of 9 mm (2 mm thick walls). To determine each well's height, a maximum volume of 120 μl was estimated (twice the 60 μl volume eventually used), which, combined with the 7 mm inner diameter, resulted to a cylinder height of approximately 3.2 mm. Moreover, the limitations on the well height transpired from the requirement for lowest possible shading. Given that the wall around each well is not transparent, the wall shading had to be considered during the design.

Each well is designed to hold a different solution: two of the wells carry the lower and upper protein concentration solutions and the third the sample under measurement. The lower protein concentration solution is in fact the blank sample, consisting of 30 μl of buffer (dist. water) and 30 μl of Bradford solution. The upper protein concentration solution consists of 30 μl of the Bradford solution and 30 μl of a sample with known protein concentration which serves as the upper detection limit.

With a min-max normalization, the lowest and highest concentrations always have standard 0 and 1 values respectively, even though the concentration range remains

adjustable according to the regions of interest. If the sample under measurement is within this range, it will be assigned a value between 0 and 1. From a known calibration curve this value corresponds to a protein concentration. If the sample has an out of range concentration, then it will be assigned either a 0 or a 1 value, with one of the lower or upper limit concentrations assuming an intermediate value, which points to the sample under measurement having a lower concentration than the lower limit or a higher concentration than the upper limit. This way, a rough qualitative estimation is still achieved.

The filament used in printing was the one closest to white already in stock at the lab. However, because the bottom of the test platform's wells was not white enough to provide a reliable and bright surface for color depiction, the paper disks used previously had to adopt this role. Furthermore, the filament and the overall structure was not waterproof, but during the experiments it was uncovered that the porosity did not cause leakages or any other noticeable effect and that the test platform could even be reused in some cases.

To create a calibration curve by using the designed test platform, the samples were prepared as shown in Table 3.5:

Sample conc.(μg/ml)	0	20	50	100	150	200	250	300	350
0.5 mg/ml BSA sol.(μl)	-	40	100	200	300	-	-	-	-
5 mg/ml BSA sol. (μl)	-	-	-	-	-	40	50	60	70
dist. water(μl)	1000	960	900	800	700	960	950	940	930

Table 3.5. *Preparing the samples for evaluation in the test platform.*

Firstly, a 6 mm diameter paper disk was carefully placed with tweezers in each test platform well. Then, 30 μl of 50% Bradford solution was placed in each well, along with 30 μl of each sample, prepared as presented in Table 3.5. The result of this process is in Figure 3.21.

Figure 3.21. *The set of three test platforms with concentrations from 0 to 350 μg/ml.*

After extracting the b* values for each well with ImageJ, the diagram in Figure 3.22 is obtained. The data points show an overall increasing trend, however the results are not conclusive about the accuracy of the measurement.

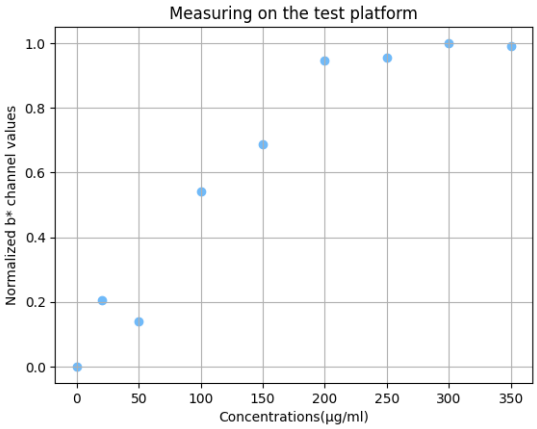


Figure 3.22. *Color extraction with ImageJ from the test platform (normalized b* values).*

In order to have a definite comparison between the calibration curve with the spectrophotometer and a calibration curve with the test platform, the following experiment was additionally performed. To compare the spectrophotometer results with the camera results, the samples used in both cases had to be of the same composition. In both measurements 50% Bradford solution was used along with a 1:1 ratio of Bradford solution to sample (BSA) solution. The original BSA solution was prepared by dissolving 20 mg of BSA into 10 ml of dist. water, resulting to a 2 mg/ml BSA solution. The samples with different protein concentrations that were eventually measured with both methods

were prepared according to Table 3.6:

Sample conc.(μg/ml)	0	25	50	75	100	150	200	250	300
2 mg/ml BSA sol.(μl)	-	25	50	75	100	150	200	250	300
dist. water(μl)	2000	1975	1950	1925	1900	1850	1800	1750	1700

Table 3.6. *Preparing the 2 ml samples for evaluation.*

The final samples under measurement were each composed of 1 ml of 50% Bradford solution and 1 ml of a different 2 ml sample (from Table 3.6). From each solution, 60 μl were measured with the phone camera on the test platform, while the rest of the volume was measured with the spectrophotometer. The data that occured from this experiment are presented in Figure 3.23.

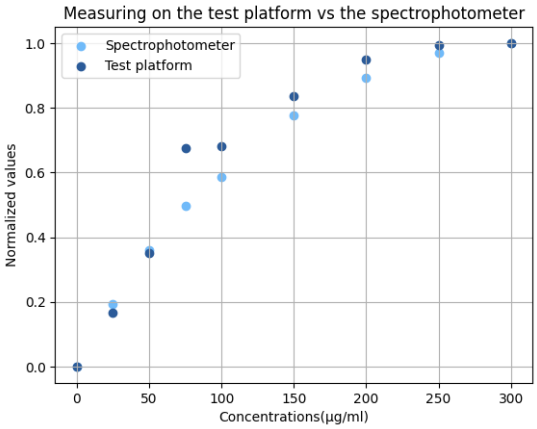


Figure 3.23. *Normalized color yields from the spectrophotometer and the test platform.*

Because the overall objective is to characterize the sample based on its protein concentration, it is useful to extract an equation that relates the normalized color value with the corresponding concentration. The curve appears to be quadratic and the fitting results to Figure 3.24. The equation is selected in a way that the estimated concentrations for 0 and 1 color are 0 and 300 μg/ml respectively. The data used for the fitting are the ones provided by normalization of the spectrophotometer results.

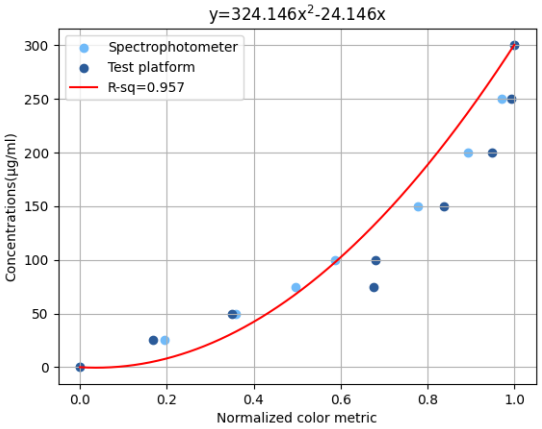


Figure 3.24. *The quadratic equation that relates the normalized color value to the sample concentration.*

Alternatively, a different fitting that still satisfies the aforementioned conditions and has a better R-squared value is presented in Figure 3.25.

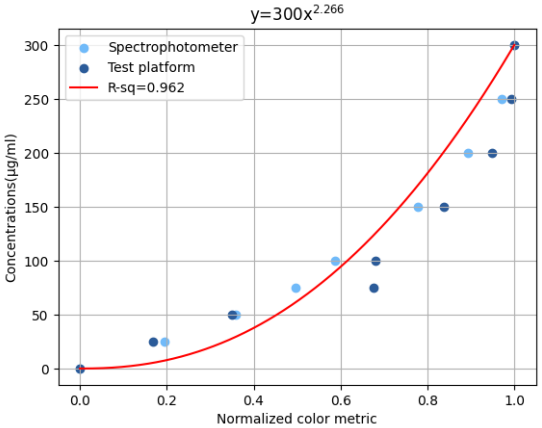


Figure 3.25. *Fitting using a power law.*

Chapter 4

Results & Discussion

There are a few remarks to be made regarding the overall procedure. Firstly, the form of the different calibration curves needs to be addressed. As also established in section 2.4.1, the relationship between absorbance and concentration is given by the Beer-Lambert law, $k = A/d = \epsilon C$, where k is the absorption coefficient, A and d are the absorbance and otical pathlengths respectively, ϵ is the molar attenuation coefficient and C is the concentration of the solution. This relationship can be applied when the analyte is significantly diluted, the intermolecular interactions are nearly nonexistent, chemical reactions are precluded, and the impact of instrumental errors is negligible. However, when one of this conditions is violated, the relationship between light absorption and concentration becomes more complicated and the linearity of the Beer-Lambert law is no longer valid [65].

In the Bradford assay this nonlinearity is introduced by the depletion of free dye in the more concentrated samples [66]. The fact that there is an overlap in the spectrum of the two color forms of the dye also contributes to this nonlinearity [67]. The linearity of the calibration curve can be improved by plotting the ratio of absorbances at 595 and 450 nm [66]. This technique was implemented by Moreira [60], who proposed that the ratio of the Blue to Green channel values of the RGB color model yields a linear response to the increase of BSA content. However, it is reported here that the B channel of the RGB model is not the ideal metric to resemble the more accurate calibration curve acquired with the spectrophotometer. For further research, a double beam spectrophotometer could be used to examine the ratio of absorbances at different wavelengths and explore the similarity with the corresponding L*a*b* model values (where b* is the channel associated with the measurement at 595 nm and a* with the measurement at 450 nm).

In literature it is repeatedly pointed out that the protein used to construct the cal-

ibration curve should be the same as the one being measured eventually [37], [66]. In many cases, nonetheless, this proves to be impractical and a common protein is used for calibration purposes. Throughout these experiments BSA has been utilized because of its readily available pure form and its inexpensiveness, qualities that contribute to its wide use for similar purposes. However, the Bradford assay response to BSA is unusually high, which might lead to underestimations when the sample under measurement contains different proteins. To compensate for this drawback, it is suggested to use a different general standard, such as bovine γ-globulin, which has a smaller dye-binding capacity, closer to the mean of other commonly found proteins [66]. An improvement on the proposed method would be to examine the response of different urine proteins to the Bradford assay and select a protein standard for the calibration curve that resembles the mean response.

Because the overall goal is to measure the protein content of urine, it is important to consider the effects of other reagents found in urine to the Bradford assay. In general, the pH of the solution influences the measurement, especially for concentrated basic buffers, where the increased absorbance might lead to overestimation [66]. Typically, for healthy individuals, there are pH fluctuations in urine in the entire physiological range, from a pH of 4.5 to a pH of 8.5 approximately [68]. The pH impact on the Bradford assay could be the objective of further research, as well as the effect of other urine substances. The presence of urea, for instance, might lead to slight underestimations [66]. Such effect was also reported here (Figure 3.12) and could be further examined by using synthetic urine as the buffer for the calibration curve and the test platform reference values.

The test platform-image analysis method proposed here is in itself an efficient, inexpensive method for the calibration of illumination conditions. Other implementations of colorimetry with smartphone require additional equipment for such a calibration, such as LEDs for constant illumination, an imaging chamber to block light variations or supplementary lenses or filters [69]. Despite the improvements in the quality of the image such equipment might provide, the cost would render them suitable only for point-of-care settings and not for individual testing at home. Other implementations of lighting calibration include normalization against a black and white background [57]. The proposed method presents an alternative form of calibration that is adjustable to a different range of concentrations according to the application. In order to reduce the complexity of the method, the samples in the wells for minimum and maximum

concentration can be substituted by reproducing the colours in a printed form. This printing can take place either by a more complex conversion of the absorbance values acquired with a spectrophotometer or with the direct use of a colorimeter.

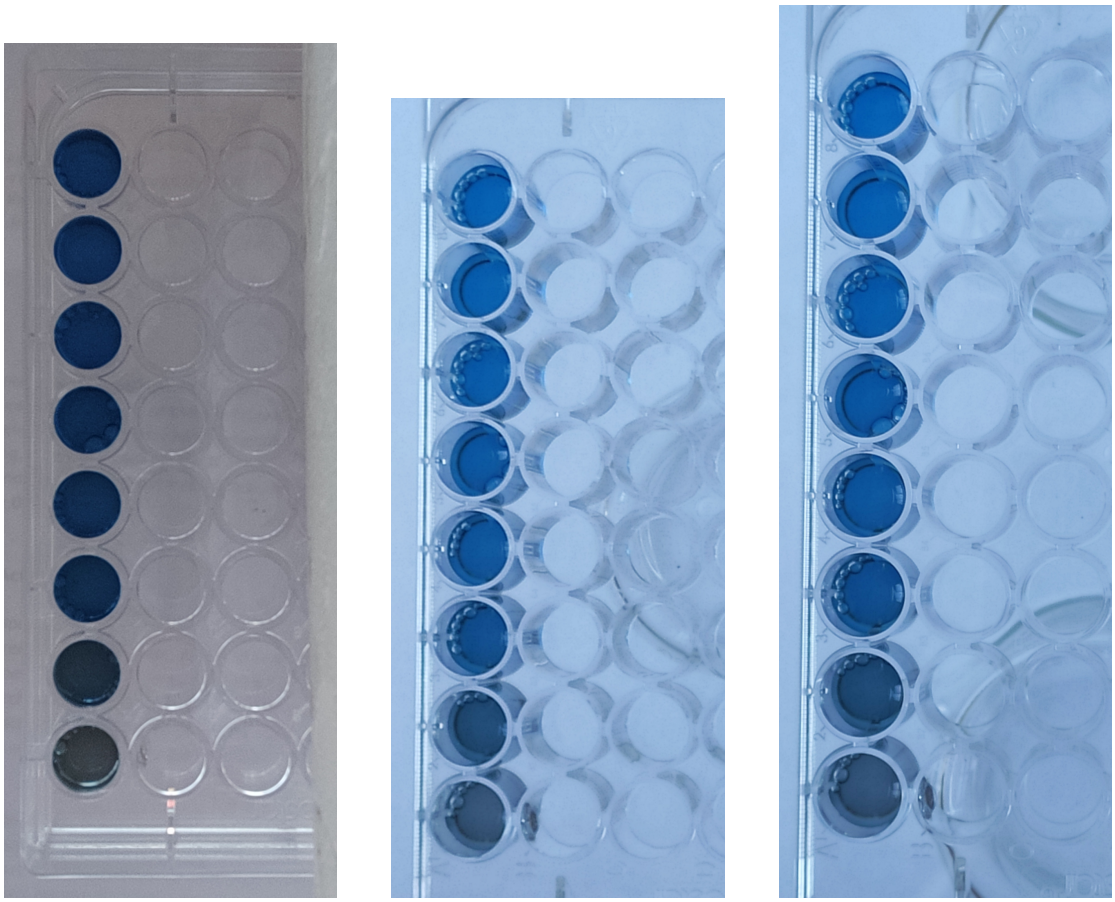
Overall, the proposed method is promising for preliminary self-testing, as it provides semi-quantitative measurements with good accuracy at a low cost. With further research it can become a significant alternative to the existing dipstick urinalysis, which has been proven an unreliable method in many cases [20], [70]. Its low cost also allows the monitoring of 24-hour urine specimen, a highly suggested alternative to the use of a single sample that can separate transient proteinuria from clinically significant cases [8]. The test platform can also be manufactured from biodegradable materials, such as polylactic acid (PLA), resulting to a more environmentally conscious solution.

Appendix A

Image analysis

A.1 Color models comparison

In this appendix supplementary images and graphs are presented concerning the various steps of image color processing. The three selected images referred to in the section First step of image color processing are shown in Figure A.1.1.



(a) *First image.* **(b)** *Second image.* **(c)** *Third image.*

Figure A.1.1. *Selected images for processing.*

It can be seen that the hue component in these pictures is the least reliable metric. As shown in figure A.1.2, the hue component remains constant despite the fluctuations of the sample concentrations in protein.

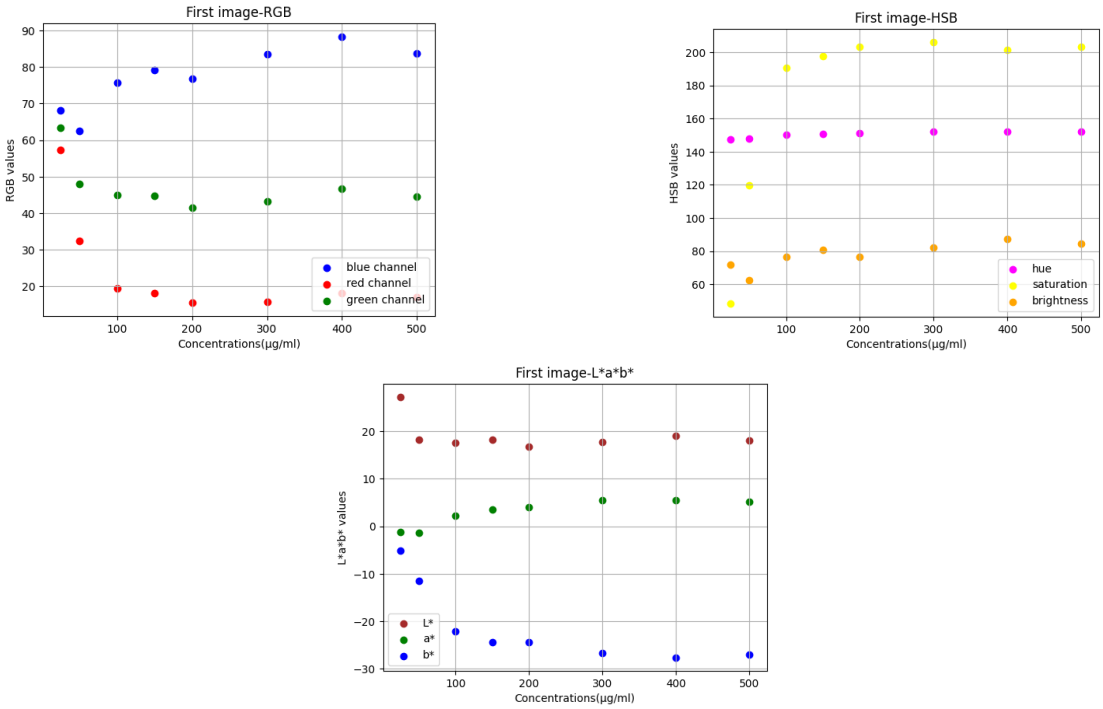


Figure A.1.2. *Channel values for the first image.*

In figures A.1.3 and A.1.4 the hue component is not depicted because during the measurement of the hue channel for the second and third image, the colored regions were not distinguishable enough from the background.

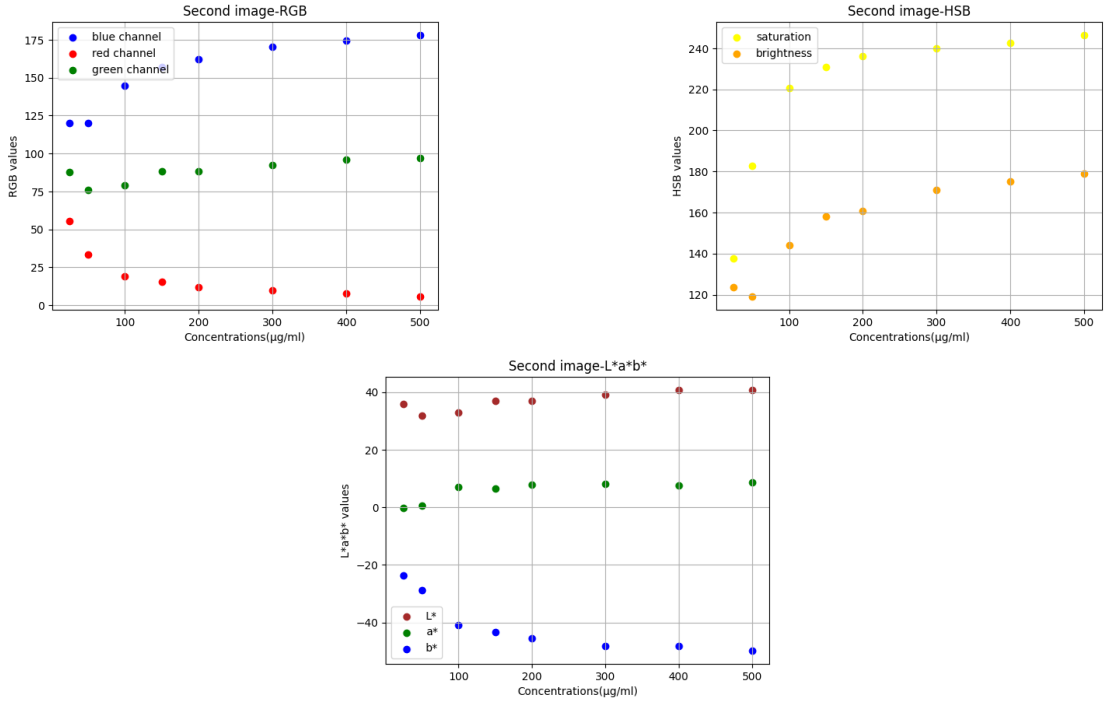


Figure A.1.3. *Channel values for the second image.*

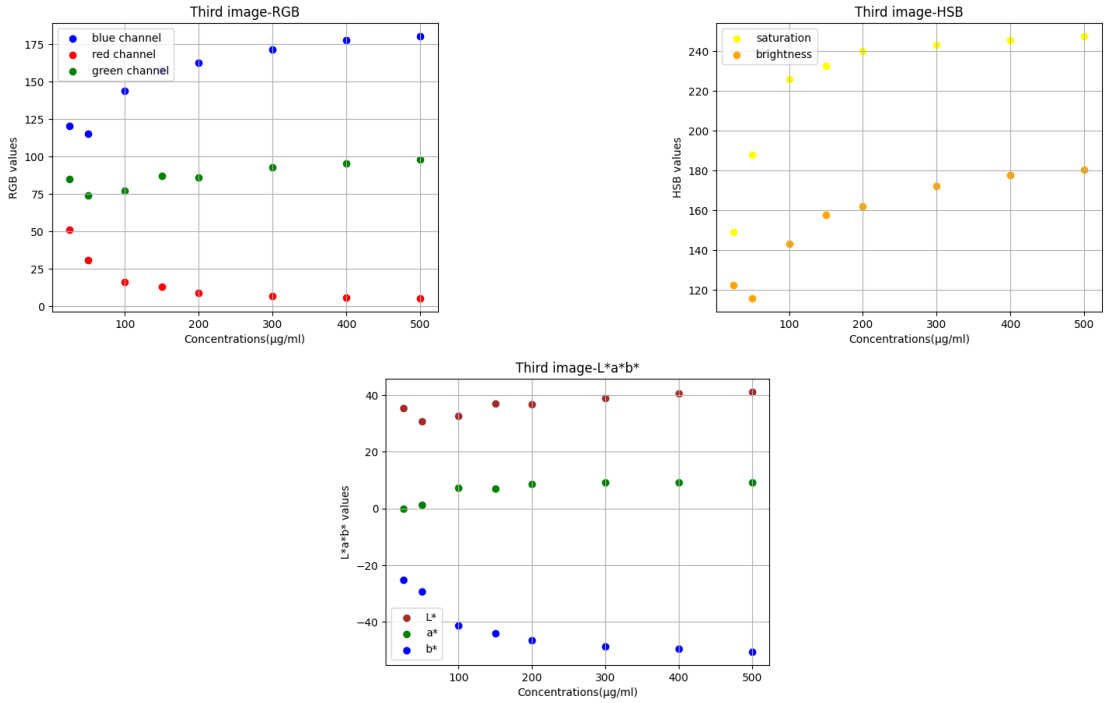


Figure A.1.4. *Channel values for the third image.*

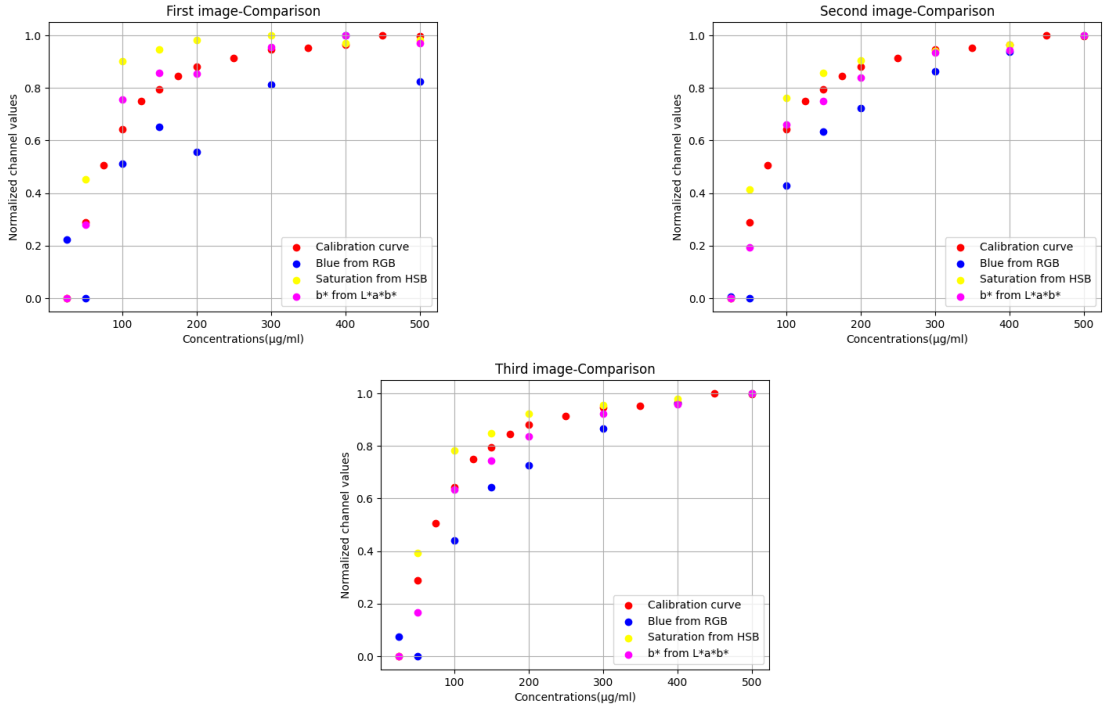


Figure A.1.5. *Proximity of the channel values to the calibration curve.*

A.2 Circle detection and color extraction

In this section I will describe the steps towards automatically extracting the b* values of the wells in the test platform (Figure 3.20) without using ImageJ. The code was developed in Python (Jupyter Notebook) using mainly the OpenCV library.

Firstly, the circles defined by the wells had to be selected from the entire image. To do so, the image was first resized by a scaling factor to a smaller size (Figure A.2.1), which helps with the selection of the parameters later on, and converted to grayscale (Figure A.2.2). A median blur filter was then applied to the grayscale image (Figure A.2.3) with a kernel size of 13. The function of this filter is shortly described in section B.2.1 of Appendix B. In our case it is used to blur the details of the image while retaining the sharper edges of the objects, to make more prominent the circles formed by the interior of the wells. The kernel size was selected via trial and error so that the circles are detected correctly in multiple images.

In the next step, the circle detection function is applied to the blurred, grayscale image, with the result of this procedure presented in Figure A.2.4. The OpenCV function used for this purpose is the HoughCircles function which uses the Hough gradient function to detect the circles. The Hough gradient function is slightly more sophisticated than the standard Hough transform, which is described in the Appendix B section B.2.3.

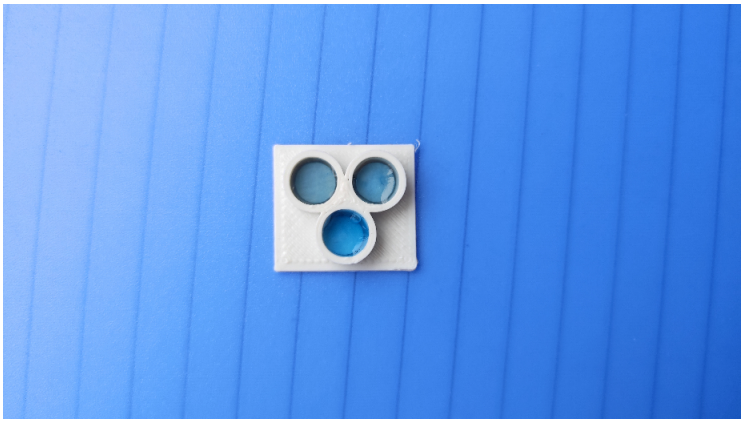


Figure A.2.1. *The resized image under processing.*

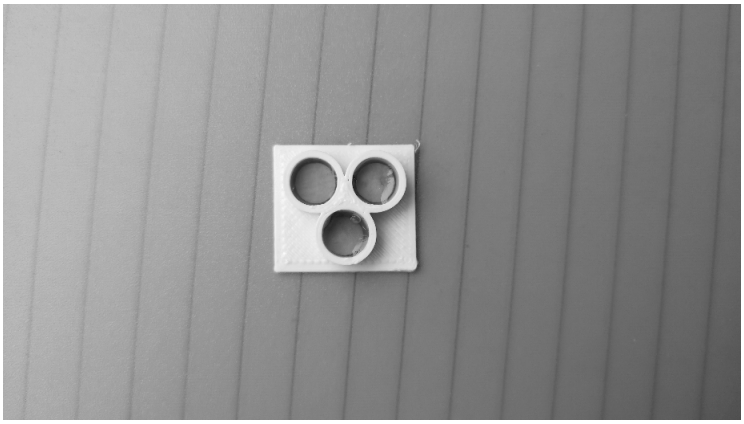


Figure A.2.2. *The image in Figure A.2.1 after grayscale transformation.*

The Hough gradient method in OpenCV works by computing the gradient for the edge pixels using the Sobel operators described in Appendix B section B.2.2. By using the gradient information, every point along the line indicated by this slope in noted in the accumulator. The centers of the circles then are selected by finding the points with the most votes in a certain neighborhood of the two-dimensional accumulator [71].

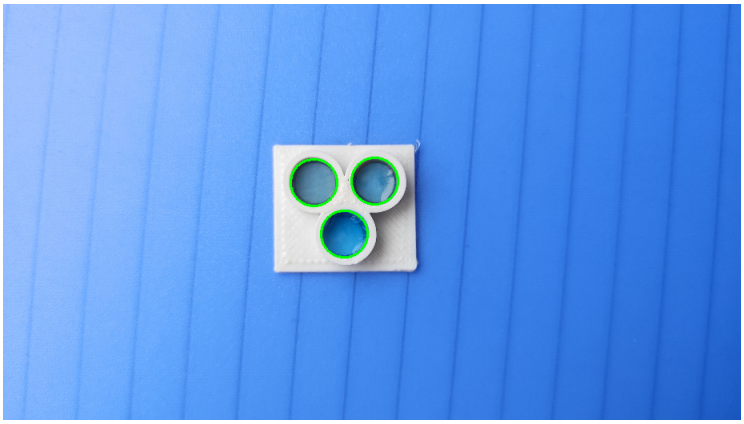


Figure A.2.4. *The detected circles drawn in the original image.*

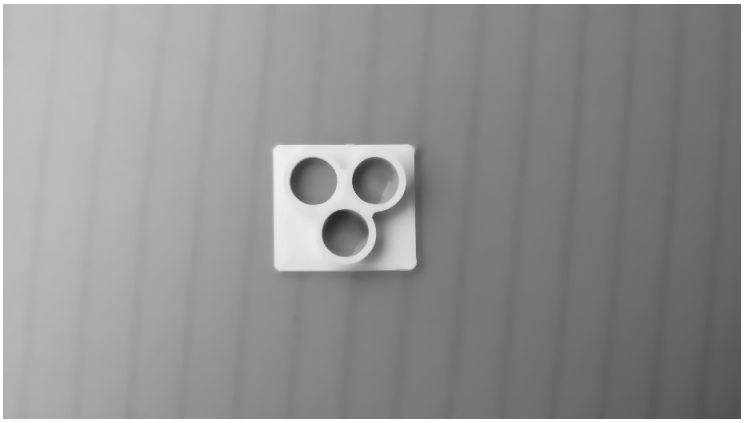


Figure A.2.3. *The image in Figure A.2.2 after median blur.*

After the successful circle detection (which needs parameter adjusting to a certain degree) the color in the inside of the circles has to be measured next. In the image however there re are some reflections and shadings that influence the result. To dull these effects gaussian filtering can be applied to the original image as shown in Figure A.2.5.

Figure A.2.5. *The image after gaussian filtering.*

In general, the shading part in the image occurs near the wall of the wells. In order to address this issue as well, the b* value can be assessed in half the radius of each well to exclude the darker values. For the color calculation, the blurred image (Figure A.2.5) was split into the three L*a*b* channels and the mean b* value of the pixels in the area defined by half the radius of each detected circle was determined. The result was printed on the surface of Figure A.2.4, as shown in Figure A.2.6.

Figure A.2.6. *Color assessment for each well by b* channel.*

A comparison of the results reached with the OpenCV implementation (after normalization) with the values obtained by manually processing the images with ImageJ (as presented in Figure 3.19) is presented in Figure A.2.7.

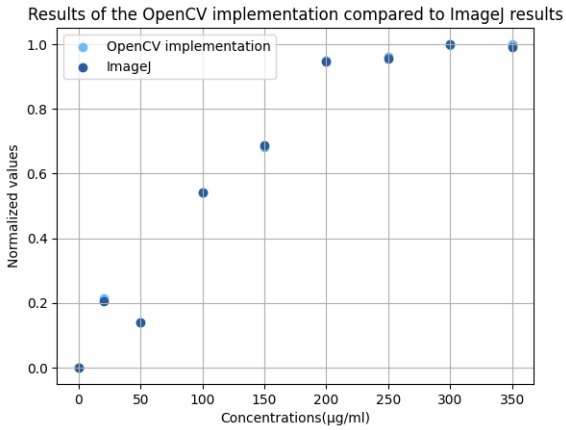


Figure A.2.7. *ImageJ analysis versus proposed OpenCV implementation.*

The implementation that was developed is therefore able to accurately extract the values of interest from an image, removing additional steps that had to be performed manually otherwise.

Appendix B

Digital Image Processing

B.1 Introductory concepts

In this appendix further information is provided on the topic of digital image processing with emphasis on color spaces, based on a review of the existing literature. The term *image* denotes a two-dimensional light-intensity function, $f(x, y)$, where the value of f returns the image's intensity or brightness at the point with spatial coordinates (x, y). Given that light is a form of energy, the function f is defined so as to receive only positive finite values, that is [72],

$$0 < f(x, y) < \infty.$$

The function $f(x, y)$ can be written as the product of two components, one representing the illumination, $i(x, y)$ and the other the reflectance, $r(x, y)$. These functions are also defined,

$$0 < i(x, y) < \infty$$

$$0 < r(x, y) < 1.$$

These equations indicate that the reflectance function is bound by total absorption (0) and total reflection (1). The illumination component is dependent on the light source characteristics, whereas the reflectance component is determined by the attributes of the objects in a scene [72].

The intensity f of a monochrome image at (x, y) can be also defined as the *gray level* of the image at that point. This value l should fall within a range

$$L_{min} \leq l \leq L_{max}.$$

The interval $[L_{min}, L_{max}]$ is the *gray scale*. This is usually shifted to an interval $[0, L]$, where $l = 0$ is regarded as black and $l = L$ as white, with all the intermediate values being shades of gray ranging from black to white [72].

A continuous image $f(x, y)$ can be reproduced as an $N \times M$ array, with elements ranging from $f(0, 0)$ to $f(N-1, M-1)$. These elements are the image's pixels. The digitization process of the image includes two separate steps of *sampling* and *quantization*. Mathematically, the sampling process can be described as a segmentation of the xy plane into a grid, with the coordinates of the center of each grid being a pair of elements from the Cartesian product $\mathbb{Z}^2$. The quantization process refers to the assignment of a gray-level value, through a function f, to each pair of coordinates (x, y) [72].

Usually, the quantities N, M as well as the number of gray levels G, are selected to be integer powers of two, so that,

$$N = 2^n, \qquad M = 2^k, \qquad G = 2^m.$$

In this case, the discrete levels are equally spaced between 0 and L in the gray scale, so that the number of bits required to store an image with these parameters would be:

$$b = N \times M \times m.$$

B.1.1 Pixel connectivity

At first, the concept of the neighbors of a pixel p should be clarified. The $4-$*neighbors* of p with coordinates (x, y) are its vertical and horizontal neighbors with coordinates

$$(x+1\ ,\ y), (x-1\ ,\ y), (x\ ,\ y+1), (x\ ,\ y-1).$$

This set of pixels is denoted by $N_4(p)$. The four diagonal neighbors of p have coordinates

$$(x+1\ ,\ y+1), (x-1\ ,\ y+1), (x+1\ ,\ y-1), (x-1\ ,\ y-1).$$

This set of pixels is denoted by $N_D(p)$. The sum of the sets $N_4(p)$ and $N_D(p)$ gives the $8-$ *neighbors* of p, denoted by $N_8(p)$ [72].

In order to define the boundaries of a specific object in an image, it is important to establish a new concept, the connectivity of pixels. The connectivity between two pixels should be determined by **(a)** whether the two pixels are somehow adjacent and

(b) whether their gray levels satisfy a certain similarity condition. Thus, for a set of gray-level values defining an arbitrary type of connectivity V (e.g. $V = 1$ for the connectivity of pixels with value 1 in a binary image), three types of connectivity can be considered:

(a) 4-connectivity: Two pixels p and q with values from V are 4-connected if q belongs in the set $N_4(p)$.

(b) 8-connectivity: Two pixels p and q with values from V are 8-connected if q belongs in the set $N_8(p)$.

(c) m-connectivity: Two pixels p and q with values from V are m-connected (mixed-connected) if:

(i) q is in $N_4(p)$, or

(ii) q is in $N_D(p)$ and the set $N_4(p) \cap N_4(q)$ is empty.

A pixel p is *adjacent* to a pixel q if they are connected in one of the three ways described above. The adjacency can be further specified in accordance to the type of connectivity. A *path* from a pixel p with coordinates (x_0, y_0) to a pixel q with coordinates (x_n, y_n) is a sequence of distinct pixels with coordinates (x_0, y_0), (x_1, y_1), ... , (x_n, y_n), where (x_i, y_i) is adjacent to (x_{i-1}, y_{i-1}), with $1 \leq i \leq n$, and n is the length of the path [72].

B.1.2 Relations and Equivalence

A *binary relation* R on a set A is a set of pairs of elements from A. If the pair (a, b) is in R, then aRb, which symbolizes that a is related to b. A binary relation R over the set A is considered:

(a) Reflexive if for each a in A, aRa.

(b) Symmetric if for each a and b in A, aRb implies bRa.

(c) Transitive if for a, b and c in A, aRb and bRc implies aRc.

Relations that satisfy all three properties are called *equivalence relations*. If R is an equivalence relation on a set A, then A can be divided into k disjoint subsets, the equivalence classes, such that aRb if and only if a and b are in the same subset [72].

These relations can be expressed in the form of a binary matrix, where the matrix elements have a value of 1 when the relation exists, or a value of 0 when the

relation is absent. For instance, for a set $A = \{a, b, c, d\}$ on which the relation $R = \{(a, a), (a, b), (a, c), (c, a), (b, d), (d, b)\}$ acts, the binary matrix B becomes:

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

If the relation R was reflexive, all the main diagonal elements of $\mathbf{B}$ would be 1, whereas if R was symmetric, $\mathbf{B}$ would be symmetric as well.

The property of transitivity as established above would imply that if a is related to b and b is related to d, as indicated by the inclusion of (a, b) and (b, d) in R, then (a, d) would also be in R. Since this is not the case, a new set can be introduced, the *transitive closure* R^+ of R, which contains exactly these relations implied by the transitivity property. In this example it would be

$$R^+ = \{(a, a), (a, b), (a, c), (a, d), (b, b), (b, d), (c, a), (c, b), (c, c), (c, d), (d, b), (d, d)\}$$

and in matrix form:

$$\mathbf{B}^+ = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

As found in [72], certain algorithmic methods exist to accurately compute the matrix $\mathbf{B}^+$ from a given binary matrix $\mathbf{B}$. The first method, proposed by Gries [73], states that

$$\mathbf{B}^+ = \mathbf{B} + \mathbf{BB} + \mathbf{BBB} + \ \ldots \ + (\mathbf{B})^n$$

where $\mathbf{B}$ is an $n \times n$ binary matrix representing a relation R over a set A of n symbols. In binary logic, the multiplications presented here are actually translated into logical ANDs and the additions into logical ORs. Applying this algorithm in the previous

example gives:

$$\mathbf{B}^{+} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

The execution of this algorithm requires on the order of n^3 AND and OR operations. A more efficient algorithm was developed by Warshall [74], using only OR operations. In this algorithm, as cited by Gonzalez and Woods in [72]:

Step 1 Set $j = 1$.

Step 2 For $i = 1, 2, \ldots, n$, if $b(i,j) = 1$, then for $k = 1, 2, \ldots, n$ set $b(i,k) = b(i,k) + b(j,k)$.

Step 3 Increment j by 1.

Step 4 If $j \leq n$, go to **Step 2**. Else go to **Step 5**.

Step 5 Stop. The result is $\mathbf{B}^{+}$ in place of $\mathbf{B}$.

Applying the algorithm into the former example gives for each iteration j the matrices (for $j = 1, 2, 3, 4$ respectively):

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

The final matrix is the $\mathbf{B}^{+}$ as calculated before.

B.1.3 Distances

The *distance function* D for the pixels p, q and z with coordinates $(x, y), (s, t)$ and (u, v) respectively should have the following properties [72]:

(a) $D(p, q) \geq 0$ ($D(p, q) = 0$ iff $p = q$),

(b) $D(p, q) = D(q, p)$,

(c) $D(p, z) \leq D(p, q) + D(q, z)$

Then, the *euclidean distance* D_e between p and q can be defined as

$$D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2},$$

the D_4 distance (or city-block distance) as

$$D_4(p,q) = |x-s| + |y-t|,$$

and the D_8 distance (or chessboard distance) as

$$D_8(p,q) = max(|x-s|, |y-t|).$$

The D_4 distance between two points is equal to the length of the shortest 4-path between these points and the same stands for the D_8 distance and the 8-path.

B.1.4 Arithmetic and logic operations

The arithmetic operations between two pixels p, q can be symbolized as follows: $p+q$ (addition), $p-q$ (subtraction), $p*q$ or pq (multiplication) and $p \div q$ (division). These operations are carried out pixel by pixel when executed for entire images [72].

The fundamental logic operations in image processing are: pANDq (or $p \cdot q$), pORq (or $p+q$) and NOTq (or $\overline{q}$). Logic operations can only be applied to binary images, while arithmetic ones to images with multivalued pixels. All these operations can also be performed on certain neighborhoods of pixels in the image, in the sense of *mask* operations. In mask operations, the value assigned to a pixel becomes a function of its gray level and the gray level of its neighbors. For instance, a very common 3×3 mask is one that when applied to a certain pixel assigns to it the following value:

$$z = w_1 z_1 + w_2 z_2 + \ldots + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

where z_i are the initial pixel values in a region 3×3 centered around z_5 and w_i are the coefficients determined by the mask, for $i = 1, 2, \ldots, 9$. By properly selecting the coefficients and applying the mask at every pixel, a variety of optimizations can be achieved, such as noise reduction, edge detection and region thinning [72].

B.2 Specialized topics

B.2.1 Smoothing filters

In this section a brief overview of the different kinds of smoothing filters in spatial domain will be presented, as summarized in Gonzalez [72]. Smoothing filters are particularly useful for noise reduction and blurring, which are very important image preprocessing steps when the goal is object extraction.

In **lowpass spatial filtering** or neighborhood averaging the n×n mask used for smoothing is:

$$R = \frac{1}{n \times n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

where R is in fact the average of all the pixels in the area of the mask and replaces the central pixel. A lowpass filter reduces or eliminates high-frequency components in the Fourier domain, while leaving low frequencies intact, resulting to image blurring [72].

When the objective is not so much to blur the image, as to reduce its noise, an optimal filtering method is **median filtering**, where the gray level of each pixel is replaced by the *median* of the gray levels in the neighborhood of that pixel. The median m of a set of values is the value for which half the values of the set are above m and the other half are below m. To perform median filtering in a neighborhood of a pixel, the values of the region should first be sorted. After this arrangement, the median value is selected and assigned to the pixel. For instance, in a 3×3 neighborhood the median is the 5th largest value, in a 5×5 neighborhood the 13th and so forth. Equal values in such a set have to be grouped during the sorting. This way, pixels with distinct intensity values in a neighborhood are isolated on the "edges" of the set and are thus forced to become more like the rest of their neighbors [72].

B.2.2 Edge detection

Edge detection is the most prevalent method for identifying significant changes in gray levels within an image. An **edge** is the boundary between two regions with fairly distinguishable gray-level properties. The edge detection is an image segmentation technique that is mainly based on the assumption that the regions are sufficiently homogeneous. This homogeneity of the regions allows for the transitions between them

to be determined by the gray-level discontinuities alone [72].

The basic concept behind most edge-detection techniques is the estimation of a local derivative operator. For instance, in the example illustrated in Gonzalez [72], a horizontal scan line is taken along a grayscale image depicting a vertical light stripe on a darker background. The profile of this line then would have a smooth change in gray level when modelling an edge (a transition from dark to light). The first derivative of this profile would be positive at the first edge, negative at the second edge and zero in regions of constant gray level, no matter whether they are dark or light. The second derivative is positive for the part of the transition related to the dark side of the edge, negative for the part of the transition related to the light side of the edge and zero in constant gray level areas. Therefore the magnitude of the first derivative can be utilized for detecting the presence of an edge in an image and the sign of the second derivative for assessing whether an edge pixel lies on the dark or the light side of an edge. The second derivative also has a zero crossing in the midpoint of the transition, a useful quality for edge detection.

The gradient of an image $f(x, y)$ at a location (x, y) is the vector:

$$\nabla \mathbf{f} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

The important quantity in edge detection is the magnitude of the gradient, usually referred to simply as the gradient of f:

$$\nabla f = \sqrt{G_x^2 + G_y^2}$$

For a simpler implementation, the gradient is approximated by the sum of the absolute values if its components:

$$\nabla f = |G_x| + |G_y|$$

An important quantity is also the direction of the gradient, which can be computed by $\tan^{-1}(\frac{G_y}{G_x})$.

In order to compute the partial derivatives at every pixel, many methods can be implemented. In particular, the method of Sobel operators has the advantage of providing both a differencing and smoothing effect, which is especially useful given that the derivatives generally enhance noise. For instance, as found in Gonzalez [72], for a 3×3

region of an image with gray levels $\begin{array}{|c|c|c|}\hline z_1 & z_2 & z_3 \\ \hline z_4 & z_5 & z_6 \\ \hline z_7 & z_8 & z_9 \\ \hline \end{array}$, the Sobel operator masks used to compute G_x and G_y at the center point are $\begin{array}{|c|c|c|}\hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$ and $\begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array}$ respectively. The derivatives then that are calculated by applying these masks at the specific region are:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

To describe the second order derivative, a useful quantity is the Laplacian, which for a 2-D function is defined as:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

The Laplacian can be digitally implemented in various ways, with the most frequent form for a 3×3 region being

$$\nabla^2 = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

or in mask form $\begin{array}{|c|c|c|}\hline 0 & -1 & 0 \\ \hline -1 & 4 & -1 \\ \hline 0 & -1 & 0 \\ \hline \end{array}$. All Laplacian masks are required to have a positive coefficient associated with the central pixel and negative coefficients related to the outer pixels. Generally the derivative implementations, as well as the Laplacian, have a zero sum of coefficients. This way, the response is zero when the mask is applied to a pixel whose value is the same with its neighbors. Although the Laplacian responds to intensity variations, it is rarely used specifically for edge detection due to its very high noise sensitivity. However, the Laplacian can be used, as previously mentioned, for determining whether a pixel is on the dark or the light side of an edge.

In general, edge detection by gradient operations is efficient in cases involving images with sharp intensity fluctuations and relatively low noise. For images with blurry edges or high noise content, a preferable technique is the zero crossings method of the Laplacian, where the image is convolved with the Laplacian of a 2D Gaussian func-

tion [72].

B.2.3 Edge linking and boundary detection

Edge detection techniques on their own are often insufficient to determine the pixels of a boundary, mainly because of the noise and non-uniform illumination in the area under examination. Therefore, boundary detection procedures are introduced after the edge detection step to formulate a meaningful boundary out of the edge pixels [72].

A very simple approach to edge linking is by analyzing the common characteristics of pixels in a small neighborhood. The similarity is determined by assessing two principal properties: **(1)** the magnitude of the gradient operator used to produce the edge pixel (∇f) and **(2)** the direction of the gradient ($\tan^{-1}(\frac{G_y}{G_x})$). The similarity between two neighboring pixels is established when the calculated (1) and (2) differ from one another by a predetermined value [72].

A widespread method for global processing, as opposed to the local method described above, is the Hough transform. If I consider a point (x_i, y_i) and the general equation of a straight line, $y_i = ax_i + b$, it is obvious that infinite lines pass through (x_i, y_i), each with a different set of (a, b). However, writing the line equation as $b = -x_i a + y_i$ yields a single line for a fixed pair (x_i, y_i). This transform leads to the **parameter space**, or *ab plane*. As illustrated in Figure B.2.1, for a second point (x_j, y_j) in the image space there is also a line in the parameter space, which intercepts the line of (x_i, y_i) at a certain point. This point, denoted (a', b') defines the image space line that contains both the (x_i, y_i) and (x_j, y_j) points. If all the points on a line in image space are transformed in lines in the parameter space, they will intersect in only one point, which gives the coefficients for a straight line in the image space [72].

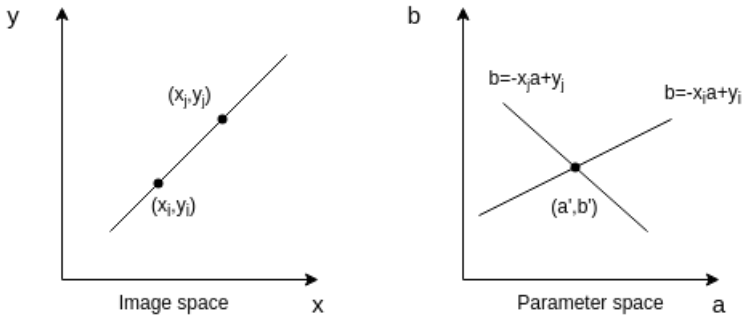


Figure B.2.1. *Hough transform from the xy plane to the parameter space.*

The algorithmic implementation of the Hough transform involves firstly the quantization of the parameter plane and the creation of an *accumulator array*. The cell of the

array in position (i,j) with a value $A(i,j)$ corresponds to the square associated with parameter space coordinates (a_i, b_j). The cells are initially set to zero. For each edge point $(x_k, y_k$ that satisfies the equation $b = -x_k a + y_k$, with the pair of a's and b's denoted as a_p and b_q, I let $A(p,q) = A(p,q) + 1$. This is essentially a voting scheme, where the most votes (the highest values) in the accumulator array correspond to intersection points in the parameter space as discussed previously.

The parameterization $y = ax + b$ for a straight line presents certain problems especially because the slope of the line can receive values $-\infty \leq a \leq \infty$, which introduces substantial memory and computational requirements for the accumulator array. A common parameterization of a line that addresses this issue is $x\cos\theta + y\sin\theta = \rho$, where both ρ and θ receive only finite values. With this transform, a straight line in the xy plane is mapped to a sinusoidal curve in the $\rho\theta$ plane.

The Hough transform is also applicable to circular shapes. For instance, if I consider the edge pixels on a circular boundary defined by the equation

$$(x - c_1)^2 + (y - c_2)^2 = r^2$$

the parameter space would be a three-dimensional space with the parameters (c_1, c_2, r), where the mapping of a point would result to a cone. This would also establish a three-dimensional accumulator array. In general, the Hough transform complexity increases with the number of coordinates and coefficients in a representation [72].

B.3 Colorimetry

In this section basic and more advanced concepts of colorimetry will be presented. Here I will mostly follow the books of Pratt [75] and Fairchild [5].

B.3.1 Certain definitions

It is important to first establish some basic terminology concerning color, color models and color appearance. Fairchild [5] uses Wyszecki's definitions to make the distinction between basic and advanced colorimetry. According to Wyszecki [76], **basic colorimetry** is employed for "making a prediction on whether two lights (visual stimuli) of different spectral power distributions will match in color for certain given conditions of observation". In order to make such a prediction, the tristimulus values of the stimuli should be specified (such a process is described in section B.3.2). **Advanced**

colorimetry broadens the purpose of the field, as its aim is to eventually assess "the appearance of color stimuli presented to the observer in complicated surroundings as they may occur in everyday life" [76]. Concepts that fall under the scope of advanced colorimetry are, for instance, the color-difference, chromatic adaptation or metamerism ("the fact that two stimuli can match in color while having disparate spectral power distributions" [5]).

The color appearance of objects exists due to the interaction of three parts: the light source, the object and the observer, the human visual system. The **light source**, as defined by Fairchild [5], is "an actual physical emitter of visible energy", with common light sources being the daylight, an incandescent lightbulb or a florescent tube. For the standardization of the color properties of light sources, the **illuminant** is introduced, "a standardized table of values that represent a spectral power distribution typical of some particular light source" [5].

According to Fairchild [5], **hue** is the "attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors: red, yellow, green, and blue, or to a combination of two of them". According to this definition, white is considered an **achromatic color**, meaning that it possesses no hue.

Brightness is the "attribute of a visual sensation according to which an area appears to emit more or less light", whereas **lightness** is "the brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting". The main difference between these two terms is that brightness is used as an absolute measure, while lightness is quite relative, accounting for changes in the illumination.

Colorfulness is the "attribute of a visual sensation according to which the perceived color of an area appears to be more or less chromatic" and **chroma** is the "colorfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears white or highly transmitting". In other words, colorfulness can be used to describe the hue intensity, and chroma is in a way the relative colorfulness. Chroma, just as lightness, remains approximately constant regardless of the change in luminance, while colorfulness increases with the luminance level. For a three-dimensional model of color perception, colorfulness or chroma define the third dimension, while hue and brightness or lightness the other two.

Saturation is the "colorfulness of an area judged in proportion to its brightness". Saturation can also be described as relative colorfulness, only this time relative to

its own brightness and not to another similarly illuminated, white appearing area. For a color stimulus with known chromaticity under constant viewing conditions, the saturation remains almost constant for all luminance levels within the range of photopic vision (see figure B.3.2), except for very high brightness.

All these definitions can also be summarized into equations [5], where for instance, as previously described, “Brightness” refers to the own stimulus’ brightness, whereas “Brightness (White)” refers to the brightness of a similarly illuminated white.

$$\text{Chroma} = \frac{\text{Colorfulness}}{\text{Brightness (White)}}$$

$$\text{Lightness} = \frac{\text{Brightness}}{\text{Brightness (White)}}$$

$$\text{Saturation} = \frac{\text{Colorfulness}}{\text{Brightness}} = \frac{\text{Chroma}}{\text{Lightness}}$$

Fairchild [5] then goes on to describe **color-order systems**, by defining their universal qualities:

1. They must be an orderly and continuous arrangement of colors.
2. They have to include a logical system of denotation.
3. They should embed perceptually meaningful dimensions.
4. They need to be complemented by stable, accurate and precise samples.

For instance, color systems that do not incorporate continuous scales or the aforementioned embodiment, like the *Pantone Color Formula Guide*, are not categorized as color-order, but as **color-naming systems** [5].

A characteristic example of a widely used color-order system is the *Munsell system*, which defines color appearance with three attributes, value, hue and chroma. Value here refers to the lightness, as defined before. The aim of the Munsell system is to define colors with uniform visual changes across each of the three perceptual dimensions. The Munsell value scale has ten main steps, where white corresponds to 10, black to 0 and grays to intermediate values, embodying linearity in the perceptual lightness and lightness difference. The hue attribute is represented by a circle, divided into five equal perceptual intervals, represented by five principal hues (yellow-5Y, red-5R, purple-5P, blue-5B and green-5G). Between these principle hues, intermediate hues are also denoted (5YR, 5RP, 5PB, 5BG, 5GY) with ten integral hues between each principle

and intermediate hue. Finally, the chroma attribute is modeled through a scale with equal visual increments from neutral samples with a chroma of zero to samples with stronger hue content with increasing chroma [5].

A **color appearance model**, as defined by the CIE Technical Committee 1-34 and found in [5], is "any model that includes predictors of at least the relative color-appearance attributes of lightness, chroma and hue". These predictors should necessarily incorporate some kind of a chromatic-adaptation transform. Fundamentally, the difference between a color model and a color space is that the first one suggests a mathematical approach towards the description of colors, while the second is a specific mapping of a color model which defines a precise range of colors, called gamut, that can be accurately represented. A color appearance model additionally deals with how colors are perceived in different viewing conditions in order to model the visual experience of color.

B.3.2 Color matching

The foundation of the trichromatic theory of color vision is the notion that any color can be created by combining specific proportions of three primary colors. In an additive color reproduction system, the three primaries are individual red, green and blue light sources that when projected on a common area, produce a colored light. In a subtractive color system, a white light is passed through a series of filters, namely a cyan, a magenta and a yellow one, to eventually reproduce a colored light [75].

Additive color matching

In an additive color matching experiment, a light beam $[C]$ with spectral energy distribution $C(\lambda)$ is projected on the surface of an ideal diffuse reflector (a surface that reflects uniformly over all directions and wavelengths). On the same surface, a reference white light $[W]$ and three primary lights $[P_1], [P_2], [P_3]$ are imaged as well. At first, the intensities of the three primaries are adjusted so that their overlapping region matches the reference white in terms of hue, brightness and saturation. Their values $A_1(W), A_2(W), A_3(W)$ are then recorded in a certain physical unit (such as watts). Consequently, the intensities of the primaries are adjusted to accomplish a match with $[C]$. If a match is achieved, the new intensities $A_1(C), A_2(C), A_3(C)$ are recorded and the *tristimulus values* occur via the

normalization:

$$T_1(C) = \frac{A_1(C)}{A_1(W)} \qquad T_2(C) = \frac{A_2(C)}{A_2(W)} \qquad T_3(C) = \frac{A_3(C)}{A_3(W)}$$

If a match cannot be achieved this way, one of the primaries, for instance $[P_3]$ is superimposed with $[C]$ and then the intensities of all three primaries are adjusted until the overlapping region of $[P_1]$ and $[P_2]$ matches the overlapping region of $[P_3]$ and $[C]$. If there is a successful match, the tristimulus values become:

$$T_1(C) = \frac{A_1(C)}{A_1(W)} \qquad T_2(C) = \frac{A_2(C)}{A_2(W)} \qquad T_3(C) = \frac{-A_3(C)}{A_3(W)}$$

Obviously, if this match does fail, the experiment is repeated with different combinations, namely $[P_1]$ with $[P_3]$ versus $[P_2]$ with $[C]$ or $[P_2]$ with $[P_3]$ versus $[P_1]$ with $[C]$. Each time the sign of the tristimulus value is negative for the primary combined with $[C]$. Finally, in the rare instance that the configurations presented so far do not provide a match, the two of the primaries are superimposed with $[C]$ and their overlapped region is put under comparison with the remaining primary. In this case, two of the tristimulus values have a negative sign, for instance:

$$T_1(C) = \frac{A_1(C)}{A_1(W)} \qquad T_2(C) = \frac{-A_2(C)}{A_2(W)} \qquad T_3(C) = \frac{-A_3(C)}{A_3(W)}$$

If a match cannot be found, the other combinations are also explored. This method is used to quantitavely specify a color. Its major drawbacks are that it is cumbersome and based on the perceptual variations of an individual observer [75].

Subtractive color matching

In a subtractive color matching experiment, an illumination source with spectral energy distribution $E(\lambda)$ is passed sequentially through three dye filters (cyan, magenta, yellow). Firstly, the dye concentrations of the filters are varied until the transmitted beam perceptually matches the reference white $[W]$. The dye concentrations are recorded as $A_1(W), A_2(W), A_3(W)$ and then varied until a match with a desired color $[C]$ is achieved. These matching values $A_1(C), A_2(C), A_3(C)$ are used to calculate the corresponding tristimulus values $T_1(C), T_2(C), T_3(C)$ as in the case of the additive color matching [75].

The additive and subtractive color systems do not have essential theoretic differences. In the subtractive system the yellow dye functions as an absorber of blue light, the magenta

dye as an absorber of green light and the cyan dye as an absorber of red light. A basis for further quantitative color measurements has been provided by Grassmann [77], who developed a set of eight axioms that define trichromatic color matching. These axioms will be presented here as found in [75].

1. Any color can be matched by a mixture of no more than three colored lights.

2. A color match at one radiance level holds over a wide range of levels.

3. Components of a mixture of colored lights cannot be resolved by the human eye.

4. The luminance of a color mixture is equal to the sum of the luminance of its components.

   For the following axioms, the use of symbol = indicates a color match, + indicates an additive color mixture and · indicates units of a color.

5. Law of addition.

$$\left.\begin{matrix}[M] = [N] \\ [P] = [Q]\end{matrix}\right\} \Rightarrow \Big([M] + [P]\Big) = \Big([N] + [Q]\Big)$$

6. Law of subtraction.

$$\left.\begin{matrix}\Big([M] + [P]\Big) = \Big([N] + [Q]\Big) \\ [P] = [Q]\end{matrix}\right\} \Rightarrow [M] = [N]$$

7. Transitive law.

$$\left.\begin{matrix}[M] = [N] \\ [N] = [P]\end{matrix}\right\} \Rightarrow [M] = [P]$$

8. Color matching.

$$C \cdot [C] = M \cdot [M] + N \cdot [N] + P \cdot [P]$$

or

$$C \cdot [C] + M \cdot [M] = N \cdot [N] + P \cdot [P]$$

or

$$C \cdot [C] + M \cdot [M] + N \cdot [N] = P \cdot [P]$$

Now that the Grassmann's color matching axioms have been established, the next step towards colorimetry is to extend the concept of the tristimulus values so that they can be obtained for any given stimulus with a known spectral power distribution $\Phi(\lambda)$.

Mathematically, the tristimulus values of the spectrum can be obtained if a unit amount of power at each wavelength λ is matched with an additive mixture of three primaries. In figure B.3.1 the set of the spectral tristimulus values for primaries at 435.6 nm (blue), 546.1 nm (green) and 700.0 nm (red) are shown. These curves are known as *color-matching functions* [5].

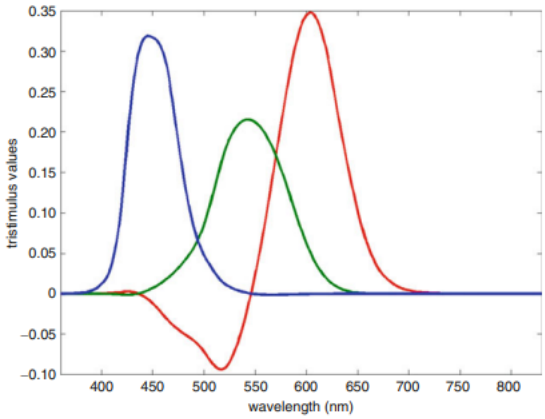


Figure B.3.1. *Spectral tristimulus values for the CIE 1931 RGB system of colorimetry with three monochromatic primaries (taken from [3]).*

Some of the spectral trimulus values shown in figure B.3.1 are negative. This shows that certain wavelengths cannot be matched using all-positive amounts of all three primaries [3]. This notion is in agreement with concepts that have already been discussed here, namely the negative tristimulus values occuring during the described additive color matching experiment and the alternative forms of Grassmann's 8th axiom.

Now that the color matching functions $\overline{r}(\lambda)$ for red, $\overline{g}(\lambda)$ for green and $\overline{b}(\lambda)$ for blue are known, the tristimulus values of a stimulus with spectral power distribution $\Phi(\lambda)$ can be calculated using the generalized equations [5]:

$$R = \int_{\lambda} \phi(\lambda)\overline{r}(\lambda)d\lambda$$
$$G = \int_{\lambda} \phi(\lambda)\overline{g}(\lambda)d\lambda$$
$$B = \int_{\lambda} \phi(\lambda)\overline{b}(\lambda)d\lambda$$

In order to eliminate the negative values in the color-matching functions, the Commission Internationale de l'Eclairage (CIE) decided to transform to another set of primaries, the XYZ primaries. Mathematically, these primaries matching all physically realizable color stimuli meant that they had to be imaginary, more saturated than monochromatic light. Moreover, the need to incorporate the CIE system of photometry

into the CIE system of colorimetry led to the selection of one color-matching function equalized to the CIE 1924 photopic luminous efficiency function [5] (figure B.3.2).

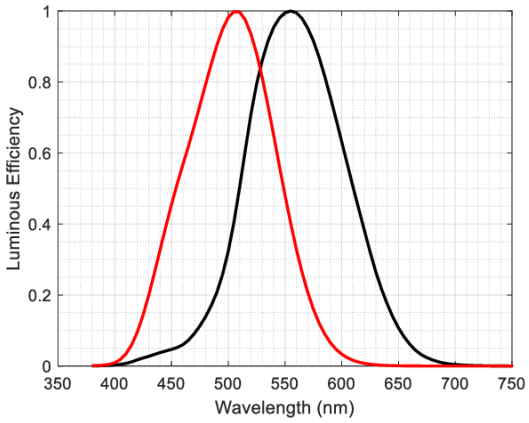


Figure B.3.2. *CIE scotopic $V'(\lambda)$ (red) and photopic $V(\lambda)$ luminous efficiency functions (taken from [4]). The function $V(\lambda)$ indicates that the visual system is more sensitive to wavelengths in the middle of the visual spectrum and becomes less sensitive when approaching the edges of the visual spectrum [5].*

Eventually, Y was selected to be the primary that would produce all the luminance response and the other two primaries, X and Z would not participate in the luminance effect. The color-matching functions for the XYZ primaries are $\overline{x}(\lambda), \overline{y}(\lambda), \overline{z}(\lambda)$ respectively, referred to as the color-matching functions of the CIE 1931 standard colorimetric observer (figure B.3.3).

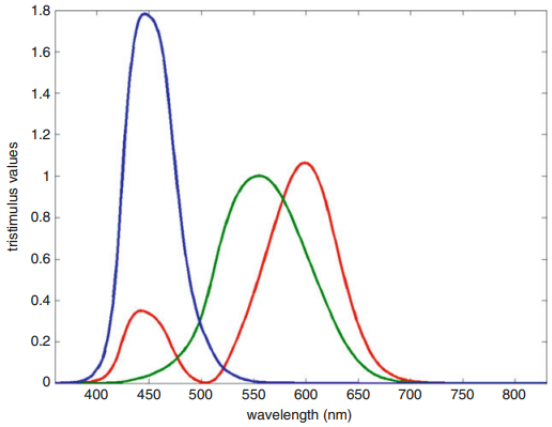


Figure B.3.3. *Color-matching functions for the CIE 1931 standard colorimetric observer (blue for Z primary, green for Y primary and red for X primary)(taken from [3]).*

The XYZ tristimulus values for a stimulus with $\Phi(\lambda)$ spectral power distribution can

be given by:

$$X = k \int_{\lambda} \phi(\lambda)\bar{x}(\lambda)d\lambda$$
$$Y = k \int_{\lambda} \phi(\lambda)\bar{y}(\lambda)d\lambda$$
$$Z = k \int_{\lambda} \phi(\lambda)\bar{z}(\lambda)d\lambda$$

where k is a normalizing constant, defined differently for relative and absolute colorimetry. In absolute colorimetry, k is 683 lumens/watt, whereas in relative colorimetry it is defined as:

$$k = \frac{100}{\int_{\lambda} S(\lambda)\bar{y}(\lambda)d\lambda}$$

With this normalization, the tristimulus values in relative colorimetry are placed between zero to 100, with Y being equal to 100 for a light source [5] or a perfectly white surface regardless of the illuminant [3].

The spectral power distribution $\Phi(\lambda)$ has different definitions, varying with the type of stimulus [5]. For instance, in the case of a reflective material with spectral reflectance $P(\lambda)$ under an illuminant of relative spectral power $E(\lambda)$ the tristimulus values become, according to [3]:

$$X = k \sum_{360}^{830} E(\lambda)\bar{x}(\lambda)P(\lambda)$$
$$Y = k \sum_{360}^{830} E(\lambda)\bar{y}(\lambda)P(\lambda)$$
$$Z = k \sum_{360}^{830} E(\lambda)\bar{z}(\lambda)P(\lambda)$$

The color-matching functions $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ were defined by the CIE at intervals of 1 nm for wavelengths in the range 360-830 nm, resulting to the former relations.

B.3.3 Chromaticity diagrams

Chromaticity diagrams were developed as a more convenient, two-dimensional representation of colors, by normalizing the tristimulus values so that the luminance information is removed. The first step towards the transformation from tristimulus values

to chromaticity coordinates is [5]:

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

From these definitions it can be derived that the sum of all three coordinates will always be equal to unity, so the third chromaticity coordinate can be written as [5]:

$$z = 1.0 - x - y$$

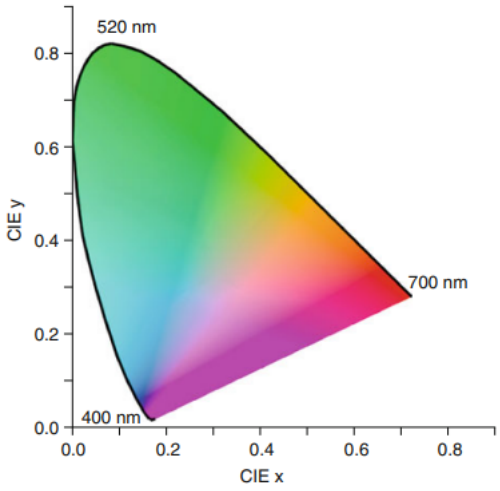


Figure B.3.4. *The 1931 CIE chromaticity diagram (taken from [3]).*

The resulting chromaticity diagram is shown in figure B.3.4. It is usually suggested that specifying the chromaticity coordinates is not sufficient to fully describe a color stimulus and to do so, one of the tristimulus values should also be reported (typically the Y value since it contains the luminance information). Then, the other two tristimulus values can be derived from [5]:

$$X = \frac{xY}{y}$$

$$Z = \frac{(1.0 - x - y)Y}{y}$$

The chromaticity coordinates are not able to provide sufficient information about the color appearance of a stimulus, since they do not account for the luminance effect or the chromatic adaptation and therefore cannot be considered perceptually uniform.

In this direction, the CIE 1976 Uniform Chromaticity Scales diagram (figure B.3.5) was developed, defined by the equations [5]:

$$u' = \frac{4X}{X + 15Y + 3Z}$$

$$v' = \frac{9Y}{X + 15Y + 3Z}$$

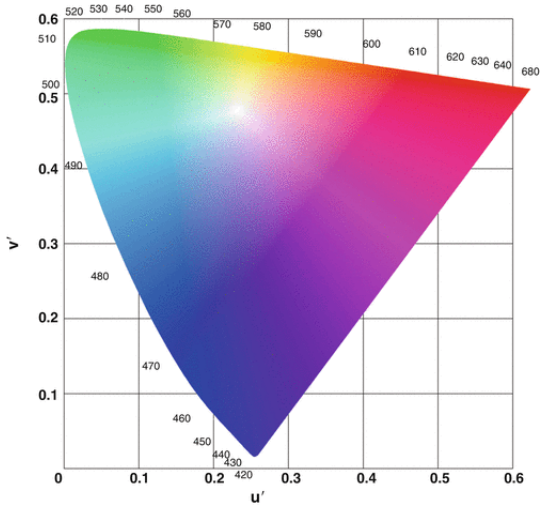


Figure B.3.5. *The 1976 CIE UCS diagram (taken from [6]).*

The 1976 UCS diagram is particularly useful for specifying the differences between colors when the primary focus is on their distinctiveness. Since it assumes that colors are of equal luminance, the UCS diagram and its coordinates are commonly employed for the representation of self-emissive colors on a screen or those produced directly from light sources [6].

B.3.4 Color spaces

In order to compensate for the lack of reliable measurement of color differences in tristimulus or chromaticity spaces, the CIELAB and CIELUV color spaces were developed by incorporating elements that account for chromatic adaptation and nonlinear visual responses [5].

The CIE 1976 (L*a*b*) color space (or CIELAB) is defined with the equations [5]:

$$L^* = 116 f(Y/Y_n) - 16$$

$$a^* = 500[f(X/X_n) - f(Y/Y_n)]$$

$$b^* = 200[f(Y/Y_n) - f(Z/Z_n)]$$

$$f(\omega) = \begin{cases} \omega^{1/3} & \omega > 0.008856 \\ 7.787\omega + 16/116 & \omega \leq 0.008856 \end{cases}$$

where two sets of CIE tristimulus values are present: those of the stimulus, XYZ, and those of the reference white, $X_nY_nZ_n$. The cube root present in this formula models the nonlinearity between the physical energy measurements and the perceptual responses [5].

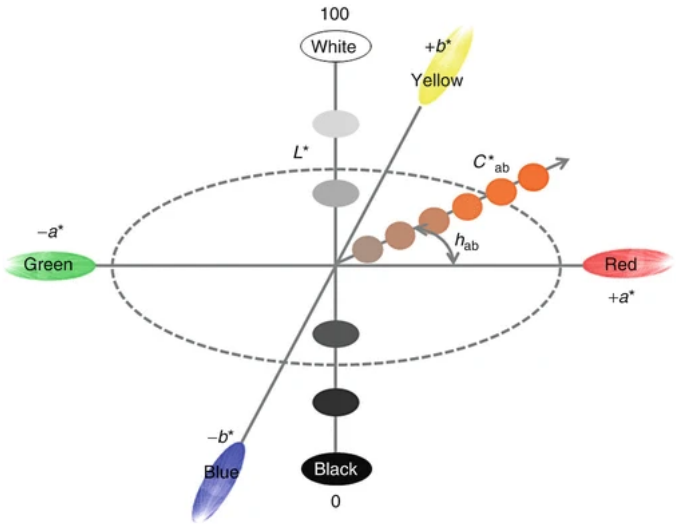


Figure B.3.6. *Schematic representation of the CIELAB color space as found in [7].*

As is illustrated in figure B.3.6, the L* measure is a correlate to perceived lightness ranging from 0.0 for black to 100.0 for white. The a* and b* values represent the red-green and yellow-blue chroma perceptions respectively, taking both negative and positive values and becoming zero for achromatic stimuli [5].

As also shown in figure B.3.6, the L*a*b* dimensions can be represented in Cartesian coordinates and can thus be transformed into cylindrical coordinates, providing predictors of chroma, C^*_{ab}, and hue, h_{ab}, determined by the following equations [5]:

$$C^*_{ab} = \sqrt{a^{*2} + b^{*2}}$$

$$h_{ab} = \tan^{-1}(b^*/a^*)$$

The lightness L*, chroma C^*_{ab} and hue h_{ab} provided by the CIELAB space substitute a simple form of a color appearance model. It is also noted that the scale L* was designed to model the Munsell system (briefly described in section B.3.1) and the Munsell

value attribute of a stimulus can be derived by dividing its CIE L* by 10. In certain applications, CIELAB performs well enough as a color appearance model, but in general it presents certain limitations, as it does not exactly match the visual physiology and does not predict luminance-dependency effects, background dependency or cognitive effects, all phenomena that should be taken into consideration when developing a color appearance model [5].

Bibliography

[1] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. Garland Science, Taylor and Francis Group, 2015.

[2] K. Barrett, S. Barman, S. Boitano, and H. Brooks, *Ganong's Review of Medical Physiology 25th Edition*. McGraw-Hill Education, 2015.

[3] S. Westland, "The cie system.," *Handbook of visual display technology*, vol. 580, pp. 161–169, 2012.

[4] R. Jenkin and C. Zhao, "Radiometry and photometry for autonomous vehicles and machines - fundamental performance limits," *Electronic Imaging*, vol. 33, pp. 211–1, 2021.

[5] M. D. Fairchild, *Color appearance models*. John Wiley & Sons, 2013.

[6] S. Westland, *USC Diagrams; Uniform Chromaticity Scales; Yu'v'*, pp. 1243–1245. New York, NY: Springer New York, 2016.

[7] V. Cheung, *Uniform Color Spaces*, pp. 187–196. Cham: Springer International Publishing, 2016.

[8] M. F. Carroll and J. L. Temte, "Proteinuria in adults: a diagnostic approach," *American family physician*, vol. 62, no. 6, pp. 1333–1340, 2000.

[9] G. Viswanathan and A. Upadhyay, "Assessment of proteinuria," *Advances in chronic kidney disease*, vol. 18, no. 4, pp. 243–248, 2011.

[10] W. Kashif, N. Siddiqi, A. P. Dincer, H. E. Dincer, and S. Hirsch, "Proteinuria: how to evaluate an important finding," *Cleve Clin J Med*, vol. 70, no. 6, pp. 535–537, 2003.

[11] M. M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," *Analytical biochemistry*, vol. 72, no. 1-2, pp. 248–254, 1976.

[12] A. Guyton, *Physiology of the Human Body*. Saunders, 1979.

[13] T. S. Larson, "Evaluation of proteinuria," in *Mayo Clinic Proceedings*, vol. 69, pp. 1154–1158, Elsevier, 1994.

[14] S. Aitekenov, A. Gaipov, and R. Bukasov, "Detection and quantification of proteins in human urine," *Talanta*, vol. 223, p. 121718, 2021.

[15] M. F. Bartal, M. D. Lindheimer, and B. M. Sibai, "Proteinuria during pregnancy: definition, pathophysiology, methodology, and clinical significance," *American journal of obstetrics and gynecology*, vol. 226, no. 2, pp. S819–S834, 2022.

[16] A. K. Leung, A. H. Wong, and S. S. Barg, "Proteinuria in children: evaluation and differential diagnosis," *American family physician*, vol. 95, no. 4, pp. 248–254, 2017.

[17] V. Verma, R. Kant, N. Sunnoqrot, and S. R. Gambert, "Proteinuria in the elderly: evaluation and management," *International urology and nephrology*, vol. 44, pp. 1745–1751, 2012.

[18] R. Machii, R. Kubota, N. Hiratsuka, K. Sugimoto, R. Masudo, Y. Kurihara, S. Kobayashi, and K. Shiba, "Urinary protein fraction in healthy subjects using cellulose acetate membrane electrophoresis followed by colloidal silver staining," *Journal of clinical laboratory analysis*, vol. 18, no. 4, pp. 231–236, 2004.

[19] J. A. Simerville, W. C. Maxted, and J. J. Pahira, "Urinalysis: a comprehensive review," *American family physician*, vol. 71, no. 6, pp. 1153–1162, 2005.

[20] E. J. Lamb, F. MacKenzie, and P. E. Stevens, "How should proteinuria be detected and measured?," *Annals of clinical biochemistry*, vol. 46, no. 3, pp. 205–217, 2009.

[21] Y.-T. Chen, H.-J. Hsu, C.-K. Hsu, C.-C. Lee, K.-H. Hsu, C.-Y. Sun, C.-Y. Chen, Y.-C. Chen, Y.-C. Yu, and I.-W. Wu, "Correlation between spot and 24h proteinuria: Derivation and validation of equation to estimate daily proteinuria," *PLoS One*, vol. 14, no. 4, p. e0214614, 2019.

[22] M. L. Champion, D. A. Becker, C. McIlwraith, C. T. Blanchard, J. M. Szychowski, D.-J. Kim, V. C. Jauk, L. M. Harper, B. M. Casey, and A. T. Tita, "Contemporary test performance of the random urine protein-to-creatinine ratio," *American Journal of Perinatology*, 2022.

[23] Y. Pasternak, D. Lifshitz, Y. Shulman, L. Hiersch, E. Rimon, M. Kuperminc, Y. Yogev, and E. Ashwal, "Diagnostic accuracy of random urinary protein-to-creatinine ratio for proteinuria in patients with suspected pre-eclampsia," *Archives of Gynecology and Obstetrics*, vol. 304, pp. 109–115, 2021.

[24] Y. Huang, X. Yang, Y. Zhang, S. Yue, X. Mei, L. Bi, W. Zhai, X. Ren, Y. Ding, S. Zhang, *et al.*, "Correlation of urine protein/creatinine ratios to 24-h urinary protein for quantitating proteinuria in children," *Pediatric Nephrology*, vol. 35, pp. 463–468, 2020.

[25] P. Zhai, Y. Huang, S. Yue, X. Yang, J. Luo, Y. Zhang, X. Mei, L. Bi, W. Zhai, X. Ren, *et al.*, "Diagnostic efficacy and influence factors of urinary protein/creatinine ratio replacing 24-h urine protein as an evaluator of proteinuria in children," *International Urology and Nephrology*, pp. 1–8, 2022.

[26] C.-F. Chen, W.-C. Yang, C.-Y. Yang, S.-Y. Li, S.-M. Ou, Y.-T. Chen, C.-J. Shih, Y.-J. Wang, C.-C. Lin, C.-C. Chien, *et al.*, "Urinary protein/creatinine ratio weighted by estimated urinary creatinine improves the accuracy of predicting daily proteinuria," *The American journal of the medical sciences*, vol. 349, no. 6, pp. 477–487, 2015.

[27] D. W. Cockcroft and H. Gault, "Prediction of creatinine clearance from serum creatinine," *Nephron*, vol. 16, no. 1, pp. 31–41, 1976.

[28] E. M. Yang, B. A. Yoon, S. W. Kim, and C. J. Kim, "Clinical utility of spot urine protein-to-creatinine ratio modified by estimated daily creatinine excretion in children," *Pediatric Nephrology*, vol. 32, pp. 1045–1051, 2017.

[29] F. Aguzzi, C. Gasparro, M. Bergami, and M. Merlini, "High-sensitivity electrophoretic method for the detection of bence jones protein and for the study of proteinuria in unconcentrated urines," *Annals of clinical biochemistry*, vol. 30, no. 3, pp. 287–292, 1993.

[30] I. A. Darwish, "Immunoassay methods and their applications in pharmaceutical analysis: basic methodology and recent advances," *International journal of biomedical science: IJBS*, vol. 2, no. 3, p. 217, 2006.

[31] R. E. Chambers, D. G. Bullock, and J. T. Whicher, "Urinary total protein estimation–fact or fiction?," *Nephron*, vol. 53, no. 1, pp. 33–36, 1989.

[32] C. M. Stoscheck, "[6] quantitation of protein," in *Methods in enzymology*, vol. 182, pp. 50–68, Elsevier, 1990.

[33] J. W. Donovan and S. Leach, "Physical principles and techniques of protein chemistry," *by SJ Leach, Academic Press, New York*, p. 101, 1969.

[34] M. H. Simonian, "Spectrophotometric determination of protein concentration," *Current Protocols in Cell Biology*, vol. 15, no. 1, pp. A–3B, 2002.

[35] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray, "How to measure and predict the molar absorption coefficient of a protein," *Protein science*, vol. 4, no. 11, pp. 2411–2423, 1995.

[36] H. Edelhoch, "Spectroscopic determination of tryptophan and tyrosine in proteins," *Biochemistry*, vol. 6, no. 7, pp. 1948–1954, 1967.

[37] J. E. Noble, "Quantification of protein concentration using uv absorbance and coomassie dyes," in *Methods in enzymology*, vol. 536, pp. 17–26, Elsevier, 2014.

[38] S. J. Compton and C. G. Jones, "Mechanism of dye response and interference in the bradford protein assay," *Analytical biochemistry*, vol. 151, no. 2, pp. 369–374, 1985.

[39] T. Marshall and K. M. Williams, "Comparison of coomassie brilliant blue protein dye-binding assays for determination of urinary protein concentration.," *Clinical chemistry*, vol. 33, no. 4, pp. 577–578, 1987.

[40] J. C. Bearden Jr, "Quantitation of submicrogram quantities of protein by an improved protein-dye binding assay," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 533, no. 2, pp. 525–529, 1978.

[41] C. W. Lim, W. N. Chisnall, Y. M. Stokes, R. Pratt, and M. J. Crooke, "Effects of sodium dodecylsulphate, dye concentration and paraprotein on coomassie blue dye-binding assays for protein in urine," *Clinical Biochemistry*, vol. 21, no. 5, pp. 277–281, 1988.

[42] M. Macart and L. Gerbaut, "Evaluation of an improved coomassie dye binding method for urinary protein assay," *Clinica chimica acta*, vol. 141, no. 1, pp. 77–84, 1984.

[43] T. Marshall and K. M. Williams, "Total protein determination in urine: elimination of a differential response between the coomassie blue and pyrogallol red protein dye-binding assays," *Clinical chemistry*, vol. 46, no. 3, pp. 392–398, 2000.

[44] J. K. Chan, J. W. Thompson, and T. A. Gill, "Quantitative determination of protamines by coomassie blue g assay," *Analytical biochemistry*, vol. 226, no. 1, pp. 191–193, 1995.

[45] T. Zor and Z. Selinger, "Linearization of the bradford protein assay increases its sensitivity: theoretical and experimental studies," *Analytical biochemistry*, vol. 236, no. 2, pp. 302–308, 1996.

[46] S. Chutipongtanate, K. Watcharatanyatip, T. Homvises, K. Jaturongkakul, and V. Thongboonkerd, "Systematic comparisons of various spectrophotometric and colorimetric methods to measure concentrations of protein, peptide and amino acid: detectable limits, linear dynamic ranges, interferences, practicality and unit costs," *Talanta*, vol. 98, pp. 123–129, 2012.

[47] M. Afkarian, M. Bhasin, S. T. Dillon, M. C. Guerrero, R. G. Nelson, W. C. Knowler, R. Thadhani, and T. A. Libermann, "Optimizing a proteomics platform for urine biomarker discovery," *Molecular & Cellular Proteomics*, vol. 9, no. 10, pp. 2195–2204, 2010.

[48] K. Schosinsky, M. Vargas, A. Luz Esquivel, and M. Chavarria, "Simple spectrophotometric determination of urinary albumin by dye-binding with use of bromphenol blue.," *Clinical chemistry*, vol. 33, no. 2, pp. 223–226, 1987.

[49] A. Waheed, K. S. Rao, and P. Gupta, "Mechanism of dye binding in the protein assay using eosin dyes," *Analytical biochemistry*, vol. 287, no. 1, pp. 73–79, 2000.

[50] K. Müller and L. Brunnberg, "Determination of plasma albumin concentration in healthy and diseased turtles: a comparison of protein electrophoresis and the bromcresol green dye-binding method," *Veterinary Clinical Pathology*, vol. 39, no. 1, pp. 79–82, 2010.

[51] J. M. McGinlay and R. Payne, "Serum albumin by dye-binding: bromocresol green or bromocresol purple? the case for conservatism," *Annals of clinical biochemistry*, vol. 25, no. 4, pp. 417–421, 1988.

[52] B. S. Antharavally, K. A. Mallia, P. Rangaraj, P. Haney, and P. A. Bell, "Quantitation of proteins using a dye–metal-based colorimetric protein assay," *Analytical biochemistry*, vol. 385, no. 2, pp. 342–345, 2009.

[53] G. B. Costa, D. D. S. Fernandes, V. E. Almeida, T. S. P. Araújo, J. P. Melo, P. H. G. D. Diniz, and G. Véras, "Digital image-based classification of biodiesel," *Talanta*, vol. 139, pp. 50–55, 2015.

[54] Y. Fan, J. Li, Y. Guo, L. Xie, and G. Zhang, "Digital image colorimetry on smartphone for chemical analysis: A review," *Measurement*, vol. 171, p. 108829, 2021.

[55] A. Garcia, M. Erenas, E. D. Marinetto, C. A. Abad, I. de Orbe-Paya, A. J. Palma, and L. F. Capitán-Vallvey, "Mobile phone platform as portable chemical analyzer," *Sensors and Actuators B: Chemical*, vol. 156, no. 1, pp. 350–359, 2011.

[56] N. Bang-iam, Y. Udnan, and P. Masawat, "Design and fabrication of artificial neural network-digital image-based colorimeter for protein assay in natural rubber latex and medical latex gloves," *Microchemical Journal*, vol. 106, pp. 270–275, 2013.

[57] J. I. Hong and B.-Y. Chang, "Development of the smartphone-based colorimetry for multi-analyte sensing arrays," *Lab on a Chip*, vol. 14, no. 10, pp. 1725–1732, 2014.

[58] M.-Y. Jia, Q.-S. Wu, H. Li, Y. Zhang, Y.-F. Guan, and L. Feng, "The calibration of cellphone camera-based colorimetric sensor array and its application in the determination of glucose in urine," *Biosensors and Bioelectronics*, vol. 74, pp. 1029–1037, 2015.

[59] A. Mathaweesansurn, N. Maneerat, and N. Choengchan, "A mobile phone-based analyzer for quantitative determination of urinary albumin using self-calibration approach," *Sensors and Actuators B: Chemical*, vol. 242, pp. 476–483, 2017.

[60] D. C. Moreira, "Rgbradford: Accurate measurement of protein concentration using a smartphone camera and the blue to green intensity ratio," *Analytical Biochemistry*, vol. 655, p. 114839, 2022.

[61] C. L. d. Camargo, M. B. Vicentini, A. L. Gobbi, D. S. Martinez, and R. S. Lima, "Smartphone for point-of-care quantification of protein by bradford assay," *Journal of the Brazilian Chemical Society*, vol. 28, pp. 689–693, 2017.

[62] H. Henderi, T. Wahyuningsih, and E. Rahwanto, "Comparison of min-max normalization and z-score normalization in the k-nearest neighbor (knn) algorithm to test the accuracy of types of breast cancer," *International Journal of Informatics and Information Systems*, vol. 4, no. 1, pp. 13–20, 2021.

[63] P. Pokhrel, S. Jha, and B. Giri, "Selection of appropriate protein assay method for a paper microfluidics platform," *Practical Laboratory Medicine*, vol. 21, p. e00166, 2020.

[64] R. D. Deegan, O. Bakajin, T. F. Dupont, G. Huber, S. R. Nagel, and T. A. Witten, "Capillary flow as the cause of ring stains from dried liquid drops," *Nature*, vol. 389, no. 6653, pp. 827–829, 1997.

[65] A. Y. Tolbin, V. E. Pushkarev, and L. G. Tomilova, "A mathematical analysis of deviations from linearity of beer's law," *Chemical Physics Letters*, vol. 706, pp. 520–524, 2018.

[66] N. J. Kruger, "The bradford method for protein quantitation," *The protein protocols handbook*, pp. 17–24, 2009.

[67] N. P. Bonjoch and P. R. Tamayo, "Protein content quantification by bradford method," in *Handbook of plant ecophysiology techniques*, pp. 283–295, Springer, 2001.

[68] V. Bilobrov, A. Chugaj, and V. Bessarabov, "Urine ph variation dynamics in healthy individuals and stone formers," *Urologia internationalis*, vol. 45, no. 6, pp. 326–331, 1990.

[69] B. Coleman, C. Coarsey, and W. Asghar, "Cell phone based colorimetric analysis for point-of-care settings," *Analyst*, vol. 144, no. 6, pp. 1935–1947, 2019.

[70] J. J. Waugh, T. J. Clark, T. G. Divakaran, K. S. Khan, and M. D. Kilby, "Accuracy of urinalysis dipstick techniques in predicting significant proteinuria in pregnancy," *Obstetrics & Gynecology*, vol. 103, no. 4, pp. 769–777, 2004.

[71] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., 2008.

[72] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Addison-Wesley Publishing Company, 1993.

[73] D. Gries, *Compiler construction for digital computers*. John Wiley & Sons, Inc., 1971.

[74] S. Warshall, "A theorem on boolean matrices," *Journal of the ACM (JACM)*, vol. 9, no. 1, pp. 11–12, 1962.

[75] W. K. Pratt, *Digital image processing: PIKS Scientific inside*, vol. 4. Wiley Online Library, 2007.

[76] G. Wyszecki, "Current developments in colorimetry," *AIC Color 73*, pp. 21–51, 1973.

[77] Grassmann, "On the theory of compound colours," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 7, no. 45, pp. 254–264, 1854.