

Using Mobile Cameras to Measure Dye Concentration in Bio-Chemistry Experiments

Jackie Fong, Andrew Lau, and Anna Velasquez

Advisor: Dr. Andrew Danowitz

External Collaborator: Dr. Amy Danowitz from Mercyhurst College, Erie, PA

June, 2015

Computer Engineering Department
California Polytechnic State University, San Luis Obispo

Table of Contents

Abstract	2
Introduction	2
Project Objectives	3
Research	3
Project/System Design	4
Results	11
Conclusion	22
Recommendations	23
References	24

Abstract

Many cutting edge chemistry experiments for high-throughput assays require measuring colored outputs from specially prepared samples. Absorbance from these samples is usually measured with a specialized piece of equipment that adds to the expense, time, and complexity of obtaining experimental results. To make these experiments more accessible, we looked at developing software to allow cell phones or consumer-grade USB cameras to measure these experiments and quantify the intensity of the output.

What we have compiled thus far is research regarding several different color spaces and attempted analysis on the many channels of color spaces such as RGB, HSV, HSL, CIE Lab, CIE Luv, XYZ, and YCrCb, that might be used to measure assay results. We concentrated on analyzing the channels and also manipulating the data to find a relationship between the computed color values and concentrations of dye in the assays. We also have created several programs that assist in extracting the channel values from the assays and to sort our results to make analysis easier.

Introduction

One of the current methods of analyzing absorbance detection of assays is to use a plate reader. This is the technique that we are trying to emulate to make absorbance detecting more accessible for people to use. The assays we are evaluating are a grid of wells with samples of solutions with varying concentrations of a specific antigen to be tested for. A solution is then added that contains an antibody for the antigen. Next, another solution is added that contains molecules that will react with the antibodies. Based off the different concentrations of the original antigen, the reaction will cause the solution in each well to change color in different shades. A darker shade indicates a higher concentration, while a lighter shade indicates a lower concentration.

To measure the concentrations of these dyed samples you need to use a plate reader. This lab equipment works by emitting a specific wavelength of light on the assay and measures how much of that light transmits through to the other side of the assay. The amount of light that passes through the solution determines the concentration of certain chemicals that do not allow light to pass through. According to Beer's law, the amount of light absorbed by the medium is proportional to the concentration of the present absorbing material or solute.

Unfortunately, equipment like this is fairly expensive and somewhat time consuming, so most smaller labs and high school classrooms might not have access to them. Because of this, our goal is to create a simple well plate analyzer that simply uses a picture taken from a cell phone, which would be cheap, fast, convenient, and efficient.

Project Objectives

The overarching goal is to find a way to analyze assays without any expensive equipment using only a cell phone and software. We aim to create an application for a phone that would assist in the classroom setting. The application would circumvent the need to buy a well plate reader for lower level courses that do not require extreme precision. The application would take care of analyzing the photo of the well plate and be able to return values containing the concentration in each well of the assay. At this point, we are aiming to return concentration information that is fairly accurate.

Our system uses software to analyze pictures taken via cell phone camera to produce dye concentration information. This enables facilities with limited resources to perform high-throughput assay analyses even without expensive equipment. For the time that we were allotted, we mainly focused on the feasibility and research of the idea.

Research

There are a few articles that involve the use of mobile phones for quantitative analysis for assays and other trays for different types of experiments. There is a mobile phone based image acquisition and quantitative ratiometric method for detecting cocaine and benzoylecgonine for biological and forensic applications, which are used on standardized rapid immunoassay strips[1]. The report discusses materials, such as the type and resolution of the camera used and different lighting settings, and methods of image analysis of rapid test strips rather than actually developing a program and/or equipment that incorporates the findings into a final design. For future development to include rapid immunoassay strips as an option to the project, this paper would be a good start.

Another project involves the integration of cell phone imaging with a microchip to detect ovarian cancer in urine[2]. This project integrates a mobile application on a cell phone that

does immediate processing of microchip ELISA results to produce colorimetric detection and reports the analyte concentration, which eliminates the use of a spectrophotometer. Although the description only discusses a general overview of the image processing method, it helped with getting an idea of how the image processing is supposed to be set up stepwise.

A plate reader system for 96-well format fluorescence assays was developed for semi-quantitative measurement results focusing on the deduction of illumination source and classification of fluorescence emitted by the transport assay[3]. The system architecture of this system is closely related to how our image processing system is set up. The HSV color space is also chosen as the best for feature extraction. The classification process used is the k-means clustering algorithm because a fixed or adaptive threshold based classification would not be sufficient for unpredictable variation in data from the variability from the living microorganisms. In a similar sense, we hope that the k-means clustering algorithm is more than sufficient for our purposes.

Lastly, a mobile phone based macroscopic image analysis system was developed for a high-throughput three dimensional cell migration assay for toxicity screening[4]. In addition to the mobile app for image and data analysis, this system requires a specific stand for the well plate and the phone to be placed, so that the imaging is consistent. The phone used also has to be an iPhone since the setup requires the camera from the phone to be in a certain position to obtain the well plate imaging. Since this project is a product that is in production, the report is not detailed in specifying what methods were used in image analysis and mainly concentrates on discussing the ring closure of the experiment.

Project/ System Design

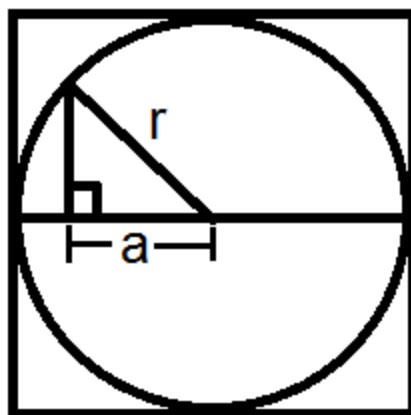
Our current design is in the form of two programs. The first program takes in a picture taken by a cell phone of the assay containing samples with various dye concentrations, automatically detects the wells in the assay, extracts the color space values of each well, and lastly outputs the found data into a specifically formatted text file. The second program will take the output of the first program and sort them into a format that corresponds to the physical positions of the assay for ease of reading and input into spreadsheet programs such as Microsoft Excel for data analysis.

PHOTO ANALYSIS

The first program takes in a file for the photograph taken as a commandline argument and uses the OpenCV library (a free open source computer vision library) in order to find an average color channel value within a well. First, it uses the OpenCV function HoughCircles to detect all the circles in the photo. Ideally, the HoughCircles function will detect all the individual wells in the assay. In practice, we have been able to detect all the wells in a picture for the majority of the time, but there was one picture where we missed detecting one of the wells. All the circles detected in the picture are stored in a vector with their X, Y coordinates and the radius.

Next, the photo is separated into three separate color channels. The three channels that the photo separates into depends on the chosen color space. As of right now, we hard code which colorspace we want to look at, but that can be changed in future implementations by adding a variable in the command line arguments. Each of the color spaces are separated into its own openCV data type Mat object, or image, in an array. With the separated channels, we can find the average of each channel per well. We use these averages per well as the data in which to correlate to the actual dye concentrations within the well.

To find the average of each well, we use the coordinates and radii stored from the HoughCircles and take a running average of a section of the well. We chose to take samples from a square of pixels within the well. During our testing, we calculated the best fit square within the well using the equation in Figure 1 with the given X, Y coordinates and radius of the corresponding detected circle.



$$r = a\sqrt{2} \quad \frac{r}{\sqrt{2}} = a \quad A = (2a)^2 = \frac{r^2}{2} \text{ for the biggest square that fits in the circle}$$

$$a = \frac{r}{2\sqrt{2}} \text{ for the square that uses } \frac{1}{2} r$$

Figure 1: Diagram and equations on how to find best fit sample square within a well

However, when printing the sampled area, we saw that with these squares, not only were we averaging the solution, we ended up including some of the white well walls. This caused our averages to skew and not match what we manually collected from the assay photo. For example, when using the RGB color space, pixels that associated with the solution tended to have high blue values and very low red and green values. The pixels associated with the well walls tended to have much higher red, green, and blue values since white reads as 255, 255, 255 in the RGB color space. The solution we took to solve this was to instead use half the radius when calculating for our sample square when averaging the wells. We chose this because when using HoughCircles, the circles found are not always guaranteed to be strictly within the wells. This helped our averages to be closer to the actual values of the pixels corresponding to the solution.

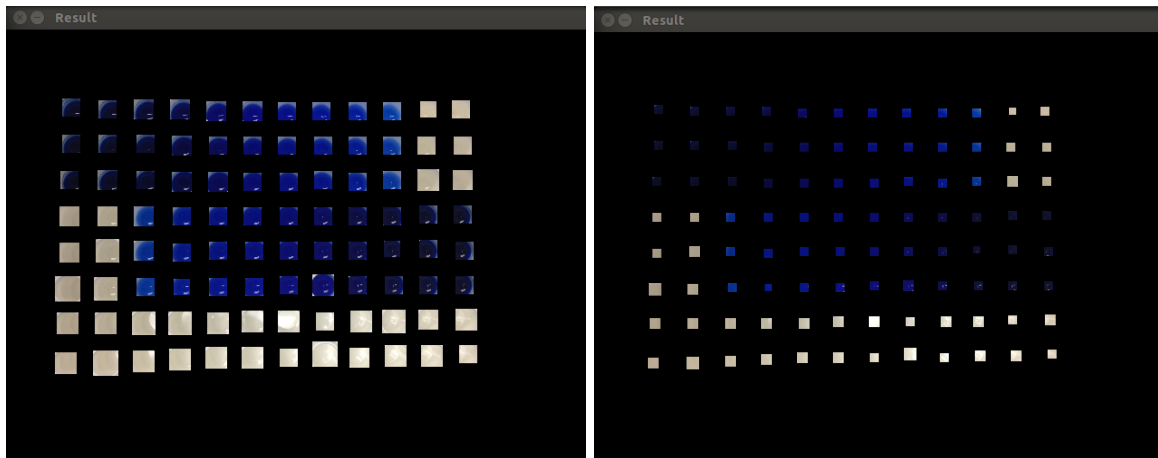


Figure 2: Best Fit Square vs. Smaller Square sample sizes

Another aspect that we found to skew our averages were the small reflections within the well solution themselves. Since our sample area became smaller than we first expected, the tiny white reflected spots started to weigh heavier in our average. To account for this, we took two averages for each well. The first time we run through the sample area, we take a regular running average with each pixel and also keep track of the highest intensity value. The second average takes into account both of these values when calculating an average closer to the real value of the solution.

We calculate the second average by running through the same sample area once more. This time, we threshold the running average by only including pixel values that are lower than halfway between the old average and the highest value. The idea behind this is that we know that the reflections will register as a light color. With white being 255, 255, 255 these “reflection” pixels are guaranteed to have a significantly higher value in all three of the RGB channels. Therefore, with this thresholding method, we should be able to skip the pixels that are reflections and our second average should more accurately represent the real values of the solution.

One limitation to note with our current implementation is that this thresholding is specific to the RGB color space. Looking at the trends of the other color spaces and comparing “solution” pixels to “reflection” pixels should determine what thresholding is needed for each color space. This would be resolved after finding the most suitable color space to use, so that it is possible to improve the accuracy of the threshold values and hard code it for that specific color space. Another option would be to first analyze the picture in the RGB colorspace in order to keep track of which pixels are “solution” pixels and which pixels are “reflection” pixels via our current method and then reject the “reflection” pixels’ values when finding the average values in the best fit color space.

After finding the average values of the wells, this program then formats and outputs the data to be parsed and sorted as seen in the format in the excerpt in Figure 3 (center: x, y radius value intensity). The outputted text file has information such as the current color space and channel; the well-specific <x,y> coordinates and values; and an end line to indicate the end of a series of wells.

Input text file for sorting:	Output of sorting:													
RGB_B	YCrCb_Cr													
center: 310, 296 19 133		1	2	3	4	5	6	7	8	9	10	11	12	
center: 270, 256 19 133		-	-	-	-	-	-	-	-	-	-	-	-	-
center: 230, 296 15 141	A	141	144	152	167	177	186	187	188	181	111	112	200	
center: 270, 296 15 138	B	138	140	143	162	174	180	182	184	179	112	112	186	
center: 270, 216 16 124	C	135	137	139	158	169	176	178	180	178	113	113	174	
center: 230, 216 16 129	D	115	114	177	177	175	169	159	149	141	139	135	25	
center: 310, 216 16 112	E	114	113	172	177	175	169	159	149	142	138	133	22	
center: 350, 216 15 95	F	114	113	179	179	179	173	164	152	144	139	136	25	
center: 270, 176 16 113	G	113	113	115	117	118	118	117	117	116	115	114	197	

Figure 3: Sample of output data text from first program and output sorted text from the second program

SORTING RESULTS

The next step we took was to make a program that would sort out all the wells and their values in a more intuitive way for data analysis. Sorting is necessary because the detected circles from HoughCircles are not stored in the same order as the assay. Therefore, to make data analysis easier, we created a program that was able to take in the previously outputted data text file via redirection. Taking in this file, we are able to parse it and sort the wells into their proper rows and columns as pictured in the actual assay. To aid us in our data analysis, we then printed the outputted data in a comma separated value format so it could be easily transferred to a spreadsheet program.

This program sorts by first finding the average radius and minimum X and Y coordinates (not coordinate pair) of the grid of wells. With this information, we now have a general knowledge of the possible range of coordinates of the first row and column. We assume that the picture taken is reasonably level. With this assumption in mind, we calculate that all wells in the first row have Y coordinate values that land in the range as depicted in Equation 1 and all the wells in the first column have X values that land in the range as depicted in Equation 2. The 25.4% leeway is to account for any minimal skewing that the well plate might have due to the picture not being perfectly level and was derived through continual testing.

$$\text{min}Y \quad \text{to} \quad \text{min}Y + 1.254 * \text{avgRad}$$

Equation 1: Range of Y coordinate values in the first row

$$\text{min}X \quad \text{to} \quad \text{min}X + 1.254 * \text{avgRad}$$

Equation 2: Range of X coordinate values in the first column

However, the values gathered for the first row and column are still unsorted within their respective row/column. For example, if rows are labeled A - E and columns are labeled 1 - 5, the first row might contain the sequence A3 A5 A2 A1 A4, while the first column might

contain the sequence 1B 1E 1A 1D 1C. It does this because HoughCircles in the first program does not detect the wells in order. To fix this, we simply run a sort on the arrays containing the first row and first column that sorts them into ascending X coordinates and ascending Y coordinates respectively. Now, with the first row and column sorted, we can use those coordinates to estimate where the rest of the wells will land.

To sort out the rest of the grid, we go through the wells one more time. We use the average well radius, X coordinate values from the first row, and the Y coordinate values from the first column to help sort the wells in the same way they were physically represented in the assay. To do this, we check each well's X and Y coordinates and compared it to the X and Y coordinates from the first row and first column respectively. If the current well coordinate falls within the first row's/column's coordinate $\pm 25.4\%$, similar to Equations 1 and 2, we consider that well to be in that row/column. The output data is then formatted in a table as shown in Figure 3 for intuitive reading. The sorted results are currently printed to screen for fast analysis, but could easily be redirected to a text file or copied over to a spreadsheet program for further breakdown.

Some minor issues with this program is that it runs on the assumption that the proper amount of wells will be given as input. If, for some reason, the first program detects more circles than what is in the assay, the sorting program will segfault because the arrays that hold the wells are hardcoded to the maximum numbers of wells in the assay. On the other hand, if for some reason the first program detects fewer circles, the sorting program will not segfault and will simply leave the spot the missing well would be as 0 concentration. However, if one of those missing wells is located in the first row or column, then the sorting algorithm will not work. The missing well affects the initial sorting of the first column/row, and therefore the subsequent sorting of the entire assay is also affected.

Refining and using these two programs enabled us to analyze our data to try and search for the best fit colorspace that would correspond to the actual dye concentrations.

COLOR SPACES

When tackling this project, the first issue we ran into was when we first tried to see how the pixel values would compare to the corresponding dye concentrations. At first, we assumed

that the blue channel in RGB would correlate linearly to the blue dyed concentrations in the wells. The higher the concentrations, the darker the solution. From initial tests, we speculated that the higher the concentration of the solution was, the higher the blue channel value would be, and vice versa for lower concentrations. However, while that conclusion was true, it turned out that there was more to take into consideration when trying to find a scale where pixel value and actual dye concentration are correlated. In addition, while we only had samples that contained blue dye, not all dye colors that are used in well plate absorbance readings will line up nicely within the RGB color space. Therefore, trying to find a color space that would be more universal to all dye colors was a priority.

The first complication we came across was when we were extracting values for the lighter concentrations. When graphing our results, we saw that the lighter wells did not change their blue value as much as the medium concentrated wells. This was because the pixels of the lighter solutions had more of the beige/cream color of the well plate instead of the light blue expected of the solution. At light concentrations, the solution was transparent, and thus the camera only picks up the color of the well. It ended up that the blue value did not change, but instead the red and green values jumped up significantly for lighter solutions.

Another complication happened when trying to extract color values on the other end of the spectrum. There came a point in the higher concentrations where the changes in the blue channel were too small and varied to be a reliable source to differentiate between the concentrations. It was then that we realized that correlating the perceived colors captured by a cell phone camera to the true colors of the solutions would be harder than first expected. Because of these two issues, we put a restriction on the range of concentration values we could measure.

Due to these limitations, we did a lot of research on the different color spaces to see if we could find a better fit between channel value and actual solution concentration value. We wanted to find a color space that had some sort of intensity or luminosity channel to detect the brightness of a photo. Our thought was that even though the pixels will not register as the true color of the solution, the differences in concentrations still go from dark to light fairly linearly. The different color spaces we looked into included HSV, HSL, CIE Lab, CIE Luv,

XYZ, and YCrCb, which are built in for the cvtColor function in openCV to convert between different color spaces.

Results

Through continual analysis of the changing data, the process and priorities have been switched many times. Our first and foremost objective was to get a color space or channel that most resembles the shape of the graph from the values given by Dr. Danowitz. These values are shown in Table 1.

WP	1	2	3	4	5	6	7	8	9	10	11	12
A	10	8	6	4	2	1	0.8	0.6	0.4	0.2	0	
B	10	8	6	4	2	1	0.8	0.6	0.4	0.2	0	
C	10	8	6	4	2	1	0.8	0.6	0.4	0.2	0	
D		0	0.2	0.4	0.6	0.8	1	2	4	6	8	10
E		0	0.2	0.4	0.6	0.8	1	2	4	6	8	10
F		0	0.2	0.4	0.6	0.8	1	2	4	6	8	10

Table 1: Table of given concentration values on the sample well plate

The first column contains labels for the rows in a well plate (WP) and the top row represents the numbers of the well plate columns. The values in between represent the given concentration values that we are trying to match. A graphical representation of the value makes it easier to see the shape we are looking for as shown in Figure 4.

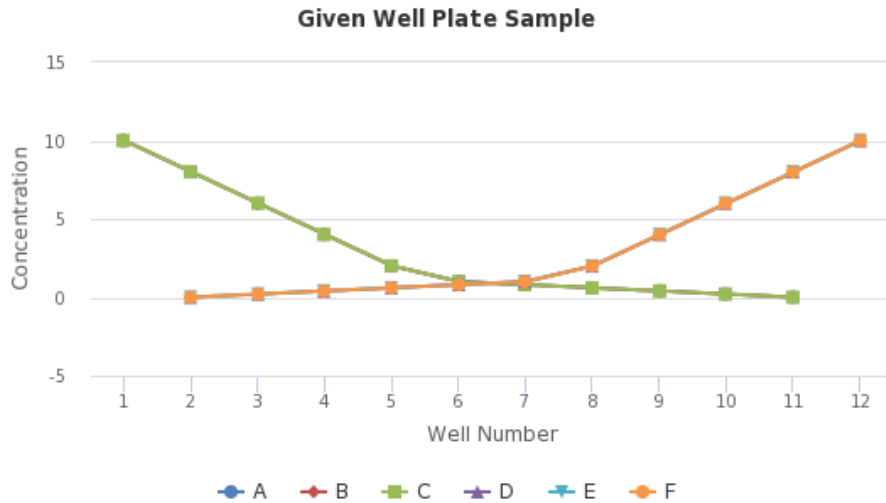


Figure 4: Ideal curve from given concentration values

The values of the concentration given to us were the same across the well plates because we wanted to make sure our color picking was reasonably accurate. Ideally, the values we obtain would track these two lines and we would be able to formulate an equation that would allow us to estimate the concentrations of the well plate after some processing. A picture of the assay that we are actually sampling can be found in Figure 12 on page 22.

Our initial approach was to look through several color spaces and pick the channels that provided a similar curve. The color spaces we attempted to extract values from are found in Table 2.

RGB	Contains channels for red, green, and blue components
HLS	Contains channels for hue, lightness, and saturation
HSV	Contains channels for hue, saturation, and value
CIE Lab (Lab)	Contains channels for lightness and color component (a, b)
CIE Luv (Luv)	Contains channels for lightness and chromaticity component (u, v)
XYZ	Contains channels analogous to the human eye LMS cone responses
YUV	Contains channels for brightness and chromaticity component (u, v)
YCrCb	Contains channels for brightness, red difference chroma (Cr), and blue difference chroma (Cb)

Table 2: Brief description of sampled color spaces

Based on our initial results, we eliminated several color spaces because the shape of their graphs did not appear to correlate with the concentration data. Figure 5 and 6 show some of the graphs we sampled for different color spaces and channels.

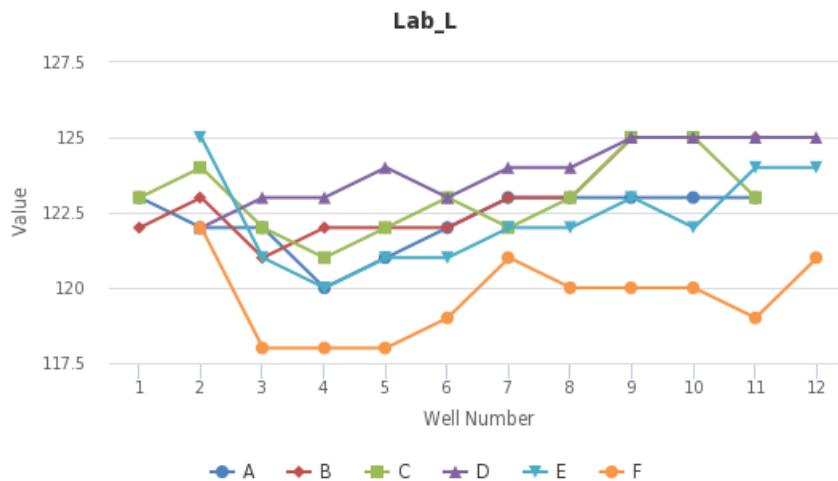


Figure 5: Lab_L does not appear to correlate with concentration data

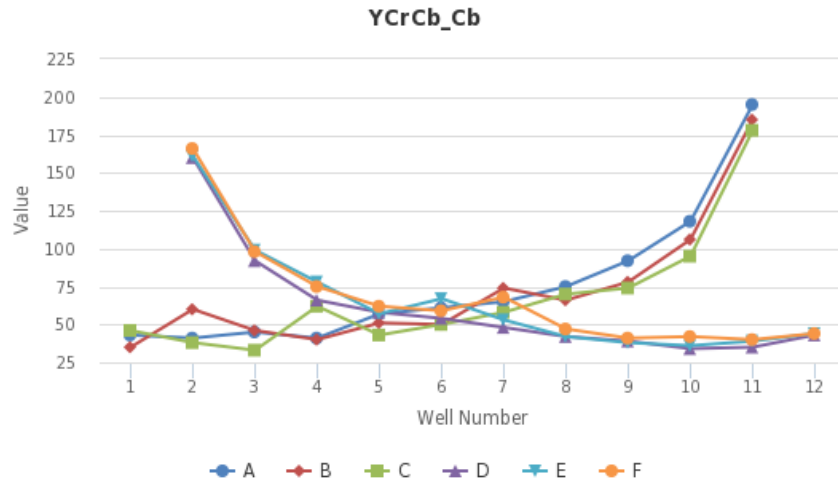


Figure 6: YCrCb_Cb appears to correlate to concentration values

We looked for shapes that would follow an overall similar curve for rows A – C and an opposite but similar curve for rows D – F. We also sampled each of the channels in those color spaces across 5 different photos of the assays and our preliminary elimination left us with the specific channels shown in Table 3.

RGB	RGB color channels contain information on how much red, green, and blue (respectively) are contained in a color.
HLS_L	The L channel in HLS contains informations of how much lightness is in a color. Low lightness means the color is closer to black, while high lightness means the color is closer to white.
HSV_H	The H channel in HSV represents hue. This is a measure of the main color as in whether a color is a type of blue, yellow, green, red, etc.
XYZ	The XYZ color space is a quantitative scale that represents color in a scale analogous to how a human eye perceives color.
CIELuv_V	The V channel in LUV represents one of the chromaticity components that make up a color. This channel along with the U channel determines where in the chromaticity diagram a color lands. Chromaticity can be described as the hue of a color in relation to a predetermined neutral white point as defined in color science.
YUV_V	The V channel in YUV represents the same channel as the V value in LUV.
YCrCb_Cb	The Cb channel in YCrCb represents an ascending scale of blue difference in chrominance.

Table 3: Brief description selected color channels

During the earlier phase of our research, we noticed that the inclusion of the well wall and their shadows within a detected circle were affecting the color values we were obtaining. As we moved further from the center of the well plate, the visibility of the well wall within a detected circle increased and therefore altered the channel values we were getting from the picture. We noticed this because several of the observed values became more erratic towards the ends of the well plate, such as data collected in columns 1-4 and 9 - 12 as seen on Figure 7. On line E, you can see a negative slope going from wells 7 to 8, but a positive slope from 9 to 10. That was an issue because the slope should be unidirectional the whole time based on our sample because within a row, the concentration is either only increasing or decreasing. This issue lead to the temporary limitation of which wells we selected to analyze, focusing on the shortened range of analyzing wells 5 - 8 specifically. We selected those wells because they are directly below the camera, giving us the most perpendicular view of the solution and thus the least amount of well wall and shadow.

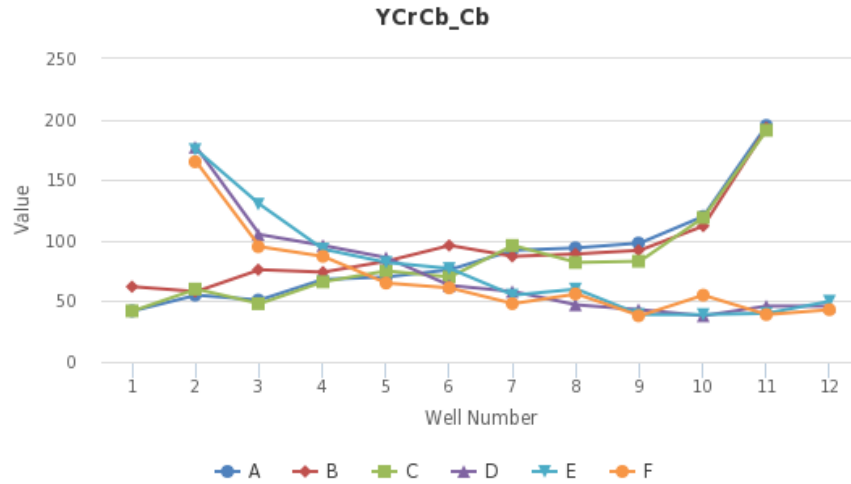


Figure 7: Slope of line F is more erratic in columns 9 - 12

The criteria for the most stable selection of values for wells 5 - 8 is that the pattern from wells 5 - 8 of rows A - C match up with the pattern from 8 - 5 of rows D - F. Furthering the elimination of channels we are studying, the channels with the best fit of this criteria turn out to be RGB_B and HSV_H which is reasonable because the increasing or decreasing blue value and the hue of the blue value should track the actual concentration of the dye accordingly. Table 4 below demonstrates that the values do follow an expected linear pattern in accordance to the graph. The first rows read of the color space and well number. Then the following rows display the averages across sets of wells and their concentrations going one way. The well rows are grouped from A-C because they have an ascending actual concentration value while well rows D-F have a decreasing actual concentration value as pictured in Figure 12. There are 4 other sets of similar tables, but they all display the same pattern with minimal differences.

RGB_B	5	6	7	8
A-C	100.7	125.7	140	141.3
D-F	134.7	125	117.3	97

HSV_H	5	6	7	8
A-C	101	126	140.3	141.3
D-F	134	125	118	97.3

Table 4: Tables of RGB_B and HSV_H's values across wells 5 - 8

Further showing that we are able to track the color, we averaged the values across the channels of Luv_v and YCrCb_b. This time, we changed our approach to sampling the photo using smaller squares and thresholding. This method proved to refine our values even further as shown in Table 5.

Actual Conc.	10	8	6	4	2	1	0.8	0.6	0.4	0.2
HSV_H Avg	41	45.3	49.7	64.3	89.5	110.5	124.2	131	139.2	152
Luv_v Avg	21.5	24	24.5	31.8	45.8	58.3	66.3	71	77.8	95.7
YCrCb_Cb Avg	24	26.2	26	29.5	36.5	43.3	48.5	52.7	60	81.7
RGB_B Avg	61.8	66.5	66.5	81.2	99.2	118.5	132	137.2	146.3	158.2

Table 5: Refined color tracking channels

We noticed that as the concentration decreases, the channel value usually increases. However, Tables 4 and 5 do not directly scale to the color value. In an attempt to establish a precise relationship between the values we are getting and the actual concentration, we used part of the Beer-Lambert's Law, and created the following ratio to compare the inverse of the values we were getting to our concentration ratio.

$$\frac{Concentration_A}{Concentration_B} : \left(\frac{Value_A}{Value_B} \right)^{-1}$$

To be more specific, we are simply comparing the ratio of the photo values to the ratios of the actual concentration values of the sample. An example calculation comparing the actual concentration of 1 and the actual concentration of 2 with the HSV_H values in Table 5 would be:

$$\frac{Concentration_1}{Concentration_2} : \left(\frac{Value_{HSVH1}}{Value_{HSVH2}} \right)^{-1}$$

$$\frac{1}{2} : \left(\frac{110.5}{89.5} \right)^{-1}$$

$$0.5 : 0.80996$$

We compared the ratio of the actual concentration to the channels that dealt with the color component in a color space. Namely, we looked at HSV_H, Luv_v, YCrCb_Cb, and RGB_B. They all provided a similar trend in their data by all having a relatively similar slope and increment. The data in Table 6 compares the actual concentration value ratios with respect to the value of 1 to the corresponding HSV_H ratios with respect to the HSV_H value that correlates to the well with the actual concentration value of 1.

The reasoning behind this is that we want to compare the ratios of the actual concentration values to the ratios of the corresponding HSV_H values. The first row in Table 6 contains the ratios of the actual given concentrations compared to the arbitrarily chosen concentration of 1 (the given concentration in well column 6 in rows A-C and well column 7 in rows D-F in our sample assay [Table 1]) so that we can apply Beer's Law and see how the corresponding HSV_H ratios compare.

For example, this means that the values in the column under 0.1 in Table 6 have an actual concentration of 10 because the actual concentration ratio is 1:10. The listed values in this column represent the ratio of the extracted color value in that specific well to the extracted color value in the well that has the actual concentration value of 1. This is why all the values in the column under 1 are also 1 since this is the arbitrarily chosen value that we base our ratios off of.

The resulting Table 6 shows that the extracted channel ratios are quite far from the actual concentration ratios. Our differences from our well value ratios to the actual concentration ratios are too far apart, and the differences are represented in Figure 8, where the extracted color value ratios are plotted against the actual concentration ratios.

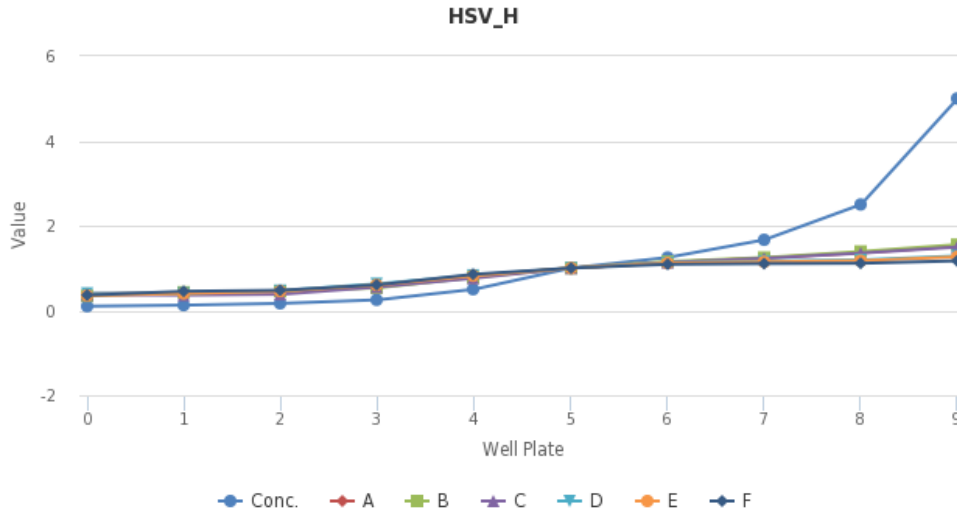


Figure 8: HSV_H value ratios plotted against the actual concentration ratio

Conc.	0.1	0.125	0.16666	0.25	0.5	1	1.25	1.6666	2.5	5
A	0.395	0.412	0.456	0.570	0.807	1.000	1.132	1.237	1.360	1.518
B	0.343	0.429	0.457	0.533	0.781	1.000	1.152	1.248	1.390	1.552
C	0.371	0.362	0.381	0.552	0.762	1.000	1.133	1.229	1.352	1.495
D	0.400	0.400	0.467	0.629	0.829	1.000	1.124	1.162	1.190	1.286
E	0.355	0.400	0.455	0.600	0.818	1.000	1.118	1.145	1.173	1.264
F	0.363	0.452	0.476	0.605	0.855	1.000	1.089	1.105	1.113	1.169

Table 6: Computed HSV values with respect of concentration ratio of 1

To reiterate, the results did not provide for a strong relationship between the extracted color values we were getting, but it allowed us to confirm that the color values we extracted from the photos were still the actual color values depicted, a shade of blue.

We then changed focus to other channels where the color component was not as relevant. We looked at the channels of HSV_S for saturation, YCrCb_Y for Luma, and CIEluv_L for light. These specific channels plotted against their concentrations provided a curve that followed the shape of the ideal curve in Figures 9 to 11 and Tables 7 to 9.

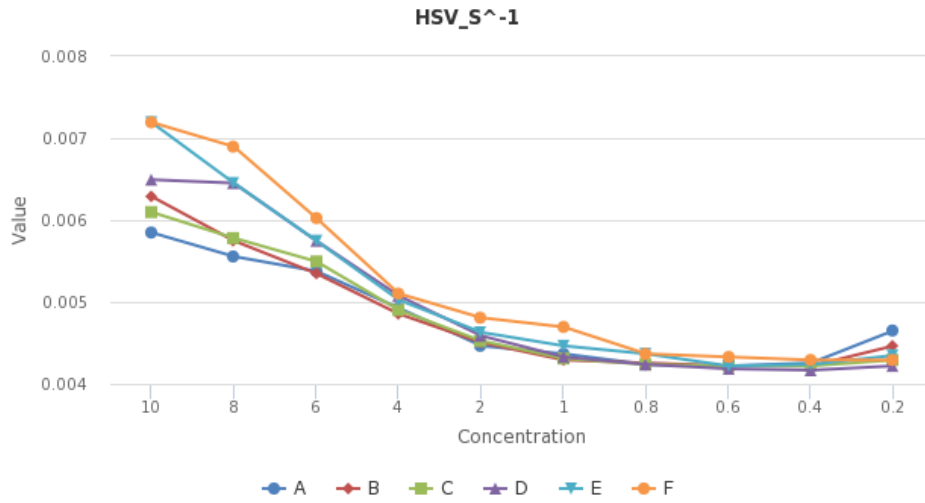


Figure 9: HSV_S channel inverted to better fit the curve

HSV_S ⁻¹	10	8	6	4	2	1	0.8	0.6	0.4	0.2
A	5.85E-3	5.56E-3	5.38E-3	4.93E-3	4.46E-3	4.37E-3	4.24E-3	4.22E-3	4.26E-3	4.65E-3
B	6.29E-3	5.75E-3	5.35E-3	4.85E-3	4.50E-3	4.29E-3	4.26E-3	4.22E-3	4.22E-3	4.46E-3
C	6.10E-3	5.78E-3	5.49E-3	4.90E-3	4.52E-3	4.31E-3	4.24E-3	4.22E-3	4.22E-3	4.29E-3
D	6.49E-3	6.45E-3	5.75E-3	5.08E-3	4.59E-3	4.33E-3	4.24E-3	4.18E-3	4.17E-3	4.22E-3
E	7.19E-3	6.45E-3	5.75E-3	5.03E-3	4.63E-3	4.46E-3	4.37E-3	4.22E-3	4.24E-3	4.35E-3
F	7.19E-3	6.90E-3	6.02E-3	5.10E-3	4.81E-3	4.69E-3	4.37E-3	4.33E-3	4.29E-3	4.29E-3

Table 7: HSV_S values inverted to fit the curve

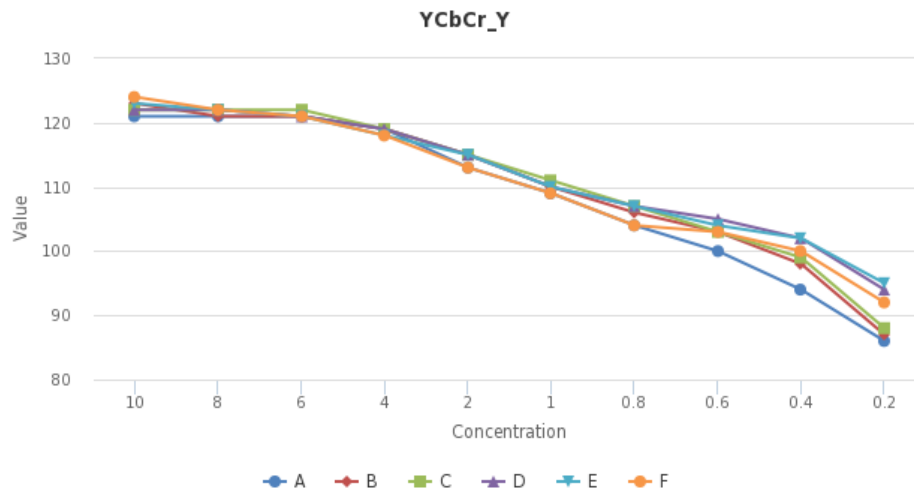


Figure 10: YCbCr_Y values with a more stable slope of value to concentration

YCrCb_Y	10	8	6	4	2	1	0.8	0.6	0.4	0.2
A	121	121	121	119	113	109	104	100	94	86
B	123	121	121	119	115	110	106	103	98	87
C	122	122	122	119	115	111	107	103	99	88
D	122	122	121	119	115	110	107	105	102	94
E	123	122	121	118	115	110	107	104	102	95
F	124	122	121	118	113	109	104	103	100	92

Table 8: YCrCr_Y values that have a more stable slope

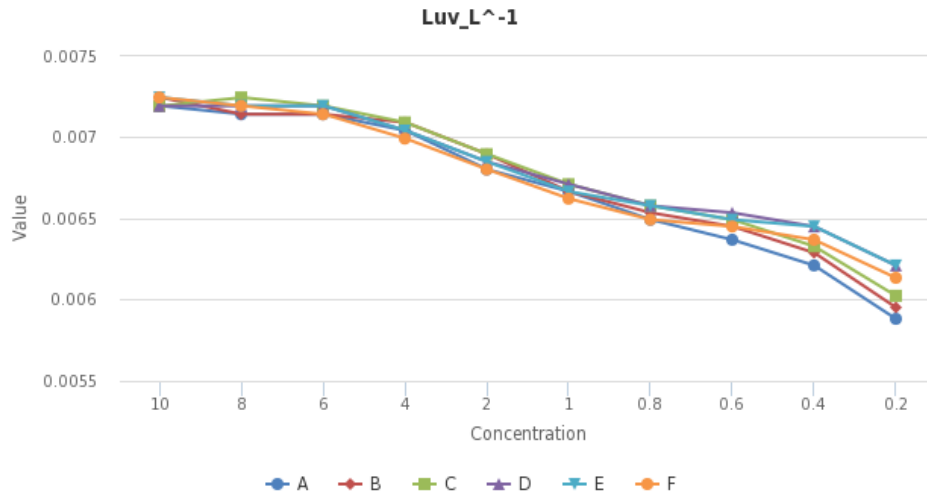


Figure 11: Luv_L channel inverted to better fit the curve

Luv_L^-1	10	8	6	4	2	1	0.8	0.6	0.4	0.2
A	7.19E-3	7.14E-3	7.14E-3	7.04E-3	6.80E-3	6.67E-3	6.49E-3	6.37E-3	6.21E-3	5.88E-3
B	7.25E-3	7.14E-3	7.14E-3	7.09E-3	6.90E-3	6.67E-3	6.54E-3	6.45E-3	6.29E-3	5.95E-3
C	7.19E-3	7.25E-3	7.19E-3	7.09E-3	6.90E-3	6.71E-3	6.58E-3	6.49E-3	6.33E-3	6.02E-3
D	7.19E-3	7.19E-3	7.19E-3	7.04E-3	6.85E-3	6.71E-3	6.58E-3	6.54E-3	6.45E-3	6.21E-3
E	7.25E-3	7.19E-3	7.19E-3	7.04E-3	6.85E-3	6.67E-3	6.58E-3	6.49E-3	6.45E-3	6.21E-3
F	7.25E-3	7.19E-3	7.14E-3	6.99E-3	6.80E-3	6.62E-3	6.49E-3	6.45E-3	6.37E-3	6.13E-3

Table 9: Luv_L values inverted to fit the curve

Note that Luv_L and HSV_S are inverted. This was done in an attempt to fit our values to the trend of the actual concentration curve in Figure 4. Additionally, the sample values in rows D - F are reversed because the actual concentrations were flipped on the well plate as seen in Figure 12. The YCrCb_Y, Luv_L and HSV_S channels seem to be the next ones to further look into for more research because they fit the curve if flipped across the y-direction. The problem would lie in scaling the numerical values into usable concentration and seeing if the relationship would follow. Once we can identify an equation or relationship, we would be able to get an estimate of the concentration through photo analysis.

Conclusion

We were not able to make a complete phone application that would be able to make estimates of the concentration in an assay, but we did a lot of research and laid down some important ground work. We have programs that are able to take in a picture and output a sorted grid of color space values. The color space analysis we have managed so far is able to detect trends that follow the actual concentration values, but we have yet to extract a reliable relationship and proper equation. We are still working on how to obtain a proper concentration through the analysis of color channels, but we think we are headed in the right direction with data analysis using the 3 most recent channels of HSV_S, YCrCb_Y, and CIELuv_L.

There are some conditions on what we would like to set for future work. The photo we analyze should be taken against a plain background on a flat surface. We would like the photo to be taken so that the assay takes up a majority of the screen as shown in the Figure 12. The photo should ideally be angled directly above the assay and reasonably levelled. Please try to not have any other miscellaneous items in the background because they may be mistaken as circles.

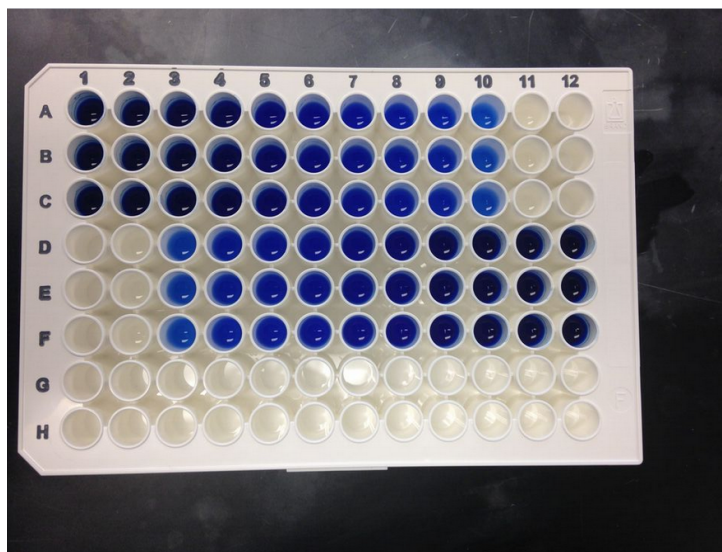


Figure 12: Example of ideal assay photo

While we focused more on finding a trend in color space values vs concentrations, some things to note is that the perceived color that a cell phone camera will pick up will never completely match the true color of the solution due to environmental and technological

factors. Factors such as different lighting, reflections, camera quality, loss of data due to image compression, and the fact that there is a definite bounded range of 8 bits of data per channel poses challenges in finding an accurate and consistent correlation between well value and actual concentration.

Recommendations

For future iterations of this project, we recommend focusing on the color space channels of HSV_S, YCrCb_Y, and CIEluv_L. However, we have only used blue dye in our experiments, so we can not guarantee that these would still be the best for other dye colors.

For the value extraction in the first program, the thresholding for separating “solution” pixels and “reflection” pixels can be further explored with k-means or mathematical thresholding. We attempted k-means thresholding, but it is not completely incorporated into the program. We were trying to use k-means with two clusters on each of the squares to separate the dye color and the assay or reflection color. After separating the pixels into clusters, we would use the separated dye pixels to get a better average, which would hopefully improve the numbers to correlate with the given dye concentrations. Hopefully these results would also help with broadening the range of the dye concentrations that can be detected.

Additionally, the current code should be rearranged so that the picture analysis program and the sorting program are one executable that takes in command line input. Command line input can include options for the path to the image file, color space, and channel you want to analyze.

Consequently, if a phone app were to be made, having a grid to guide the user when taking a photo would help yield consistent images for analysis. It may also be advantageous to code for optional different dye colors and assay sizes.

References

- 1) Cadle, Brian A., Kristin C. Rasmus, Juan A. Varela, Leah S. Leverich, Casey E. O'Neill, Ryan K. Bachtell, and Donald C. Cooper. "Cellular Phone-Based Image Acquisition and Quantitative Ratiometric Method for Detecting Cocaine and Benzoylcegonine for Biological and Forensic Applications." *National Center for Biotechnology Information*. U.S. National Library of Medicine, Apr. 2010. Web.
- 2) Wang, ShuQi, Xiaohu Zhao, Imran Khimj, Ragip Akbas, Weiliang Qiu, Dale Edwards, Daniel W. Cramer, Bin Ye, and Utkan Demirci. "Result Filters." *National Center for Biotechnology Information*. U.S. National Library of Medicine, 21 Oct. 2011. Web.
- 3) Durai Arun P, Sankaran K, Muttan S. "Image Analysis System for 96-well Plate Fluorescence Assays." *IEEE Xplore*. IEEE, July 2012. Web.
- 4) Timm, David M., Jianbo Chen, David Sing, Jacob A. Gage, William L. Haisler, Shane K. Neeley, Robert M. Raphael, Mehdi Dehghani, Kevin P. Rosenblatt, T. C. Killian, Hubert Tseng, and Glauco R. Souza. "A High-throughput Three-dimensional Cell Migration Assay for Toxicity Screening with Mobile Device-based Macroscopic Image Analysis." *Scientific Reports Sci. Rep.* 3 (2013): n. pag. *Nature.com*. Nature Publishing Group, 21 Oct. 2013. Web.