# HOW CAN A GOOD-QUALITY 3D IMAGE BE EFFICIENTLY OBTAINED FROM MULTIPLE 2D IMAGES?



Ritwik Dutta

7A #2

Stratford Middle School


Category: Science and Engineering

# Table of Contents

# Abstract

3D movies were once popular in the 1950s, but then they stopped coming out for many years till they started appearing again in the 1980s. Up until recently, 3D media has not been in the mainstream, restricted mostly to fun and hobby. The situation, however, has changed drastically in the last one year. 3D TVs, 3D Blu-ray players, and special 2-lens 3D cameras are the latest, greatest, and hottest in the world of consumer entertainment today; 3D images abound on the internet, and every major motion picture is being filmed in 3D.

A 3D movie or a 3D picture yields such a realistic impression of depth that the viewer feels he/she is present at the scene itself! This offers a more immersive experience than the regular 2D media we have been traditionally used to.

3D viewing has its roots in *stereopsis, i.e.,* depth perception that occurs naturally in humans, when each eye sees a slightly different view of the world (by virtue of the different eye positions on our faces), and the brain combines the different views yielding an impression of depth.

It follows that if we can construct an image made up of two color layers that are horizontally offset with respect to each other, we can produce a stereoscopic 3D effect when viewed through glasses with complementary color filters, such

that each eye sees its intended image. These color-based 3D images are called *anaglyphs*.

In this project, I show that using any standard digital camera and a distance-measuring apparatus mounted on a stand where the camera can be positioned and moved laterally, we can shoot two pictures of a scene by shifting the camera position horizontally, and the images can be *easily* combined to generate an anaglyph. The quality of 3D perception of such an anaglyph image depends upon the color filters chosen for viewing, the amount of lateral shift of the camera, and the image processing techniques employed. Based on my experiments, I have identified that the red-cyan color filters produce the best anaglyphs, the ideal separation is approximately the distance between the human eyes, and the red pixels need to be give an intensity boost for *good-quality* 3D perception.

## Question, Hypothesis, and Purpose

**QUESTION:** How can *good-quality* 3D images be *efficiently* obtained from 2D images?

**HYPOTHESIS:** If two regular (2D) color photos of any scene, taken by a standard digital camera by shifting the camera horizontally, can be superimposed and viewed through complementary color glasses, then it will enable 3D perception. If the images are photographed with a camera shift approximately equal to the distance between the human eyes and the images are superposed for viewing through red-cyan glasses, with the red pixels given an intensity boost to compensate for the light loss through the red filter, the overall experience will be the best.

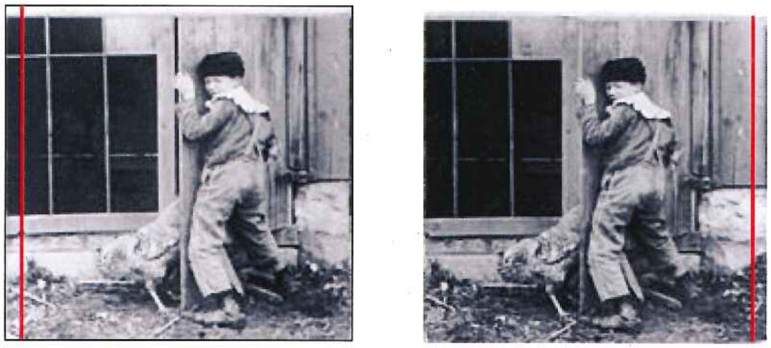
**PURPOSE:** A 3D image conveys more information than 2D. The purpose of my project is to determine how 3D images (anaglyphs) can be generated in a cost-effective manner and how these images can be viewed comfortably without eye strain. The techniques identified can be used by professionals (television studios, film producers, medical examination, product design, *etc.*) as well as amateurs (home photo, Youtube content, *etc.*) to generate and distribute optimal 3D content for regular consumption.

## Research

Perception of depth is the ability to calculate the relative position of an object based on the viewer's current position. Depth is what makes an object look three-dimensional. Without depth perception, humans would not be able to judge how far away an object is, and our eyesight would not be very good.

There are several different depth cues that are used by the brain to achieve 3D perception. These depth cues can be monocular (one eye) or binocular (two eyes). Examples of monocular cues are how objects overlap, relative sizes of objects, relative motion, *etc*. One of the most important binocular cues is *parallax*, where each eye sees a slightly different view of the world. Because the two eyes are located at different positions on the head, they see different views; the human brain combines these two views --- the left eye view and the right eye view --- and uses the differences between the images seen by the eyes to obtain and render depth information.

Unlike horses or cows, humans have two eyes located side-by-side in the front of their heads. Because of the side-by-side positioning, each eye sees the same scene from a slightly different angle. The two eye views are mostly common, but because of the angular differences each eye picks up a little bit of visual information the other does not. As shown in Figure 1, the left eye sees a little more information to the left (of the red line in Figure 1.a) and the right eye sees a little more information to the right (of the red line in Figure 1.b).



(a) (b)

**Figure 1. Left (a) and Right (b) eye views**

Therefore, each eye captures its own view and the two separate images (from the two eyes) are processed by the brain to produce one "united" picture (Figure 2). The combined image conveys more information than the constituents ... it conveys depth and is thus a three-dimensional stereo picture.

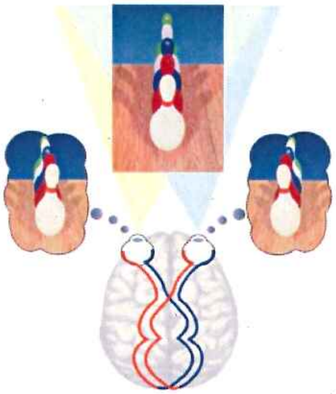


**Figure 2. Illustration of *stereopsis* where the brain fuses left and right images**

It follows from above that if the left-eye and the right-eye views of an image or a scene are presented to the eyes in a way that each eye sees only its own view, then one will be able to perceive depth and see 3D. The two eye-views can be obtained by taking two photos of an object or a scene with horizontally-shifted camera positions. Each eye can then be made to see its intended image using the idea of "complementary colors".

Complementary colors are those that are mixed to produce a neutral color (white or black). So, if the complementary color layers of a left-right image pair can be superimposed to obtain a final image --- called an *anaglyph* --- the image can then be viewed through a pair of complementary color filters to perceive 3D. This is because when the human eye sees the 3D image through the filter, it cannot see the color layer on the image that complements the color on the filter in front of the eye. This allows each eye to see a different (*i.e.,* its own) image. Some examples of complementary color pairs are red-cyan, amber-blue, and magenta-green.

To generate an anaglyph from the left and the right images, one must take some parts of the left image (L), and some parts of the right image (R) and combine them to form the final image. All digital images are composed of three layers: red (R), green (G), and blue (B); each layer is a separate image with only one color assigned to the layer (red, green, or blue).

The red-cyan glasses have the red filter in front of the left eye, so the red component of the final anaglyph image must be taken from the left image. For a similar reason, the blue and green (cyan) components of the final image must come from the right image.

Let ($R_L$, $G_L$, $B_L$) be the color planes of the left image and ($R_R$, $G_R$, $B_R$) be the color planes of the right image. Therefore, to generate a red-cyan anaglyph, one must make the final image using $R_L$, $G_R$ and $B_R$. Color-Code glasses have amber (red plus green) on the left, and blue on the right; so, to generate a Color-Code anaglyph, the color planes of the final image should correspond to $R_L$, $G_L$, and $B_R$.

In the real world, when we see different scenes or objects, both eyes receive an equal amount of light from the scene or the object. When viewing images through the anaglyph glasses, therefore, it is important that the same principle is maintained. However, different color filters may allow different amounts of light to go through. This will cause a misbalance between the eyes (in the amount of light received), hence eye strain after prolonged use. As a result, it is necessary to boost the intensity of the pixels for the color layer(s) of the image corresponding to the filter that is darker (*i.e.,* allows less light to go through).

This project necessitates a method to score data based on the overall quality of 3D perceived. Quality measurements are always subjective. So, to introduce some objectivity into it, the quality part of the *ITU BT.500 five-point quality and impairment scale* (Table 1) is used. This scale offers a standard in consumer quality evaluation.

**Table 1. ITU BT.500 five-point quality and impairment scale**

| Five-grade scale | | | |
|---|---|---|---|
| Quality | | Impairment | |
| 5 | Excellent | 5 | Imperceptible |
| 4 | Good | 4 | Perceptible, but not annoying |
| 3 | Fair | 3 | Slightly annoying |
| 2 | Poor | 2 | Annoying |
| 1 | Bad | 1 | Very annoying |

# Variables, Materials and Method

## VARIABLES

Control group: Regular (2D) pictures taken by the camera by placing it on the platform built to take 3D pictures

Independent variables: (a) Color filters chosen, and (b) the level of separation

Dependent variable: Overall 3D perception (score)

## MATERIALS

- A digital camera that saves images in the JPEG format
- A camera platform with an attached ruler that allows moving the camera horizontally by desired distances
- Subject(s) to photograph
- An NXT brick and light sensor
- A narrow-beam LED flashlight
- Red-cyan and Amber-blue (Color-Code) glasses
- Survey sheets
- 20 human subjects to get the survey
- A computer with
  - Python 2.6.6 programming language
  - PIL (Python Imaging Library) 1.1.7 patched
  - My program for image processing and anaglyph

generation
- MS Office (for producing graphs and report)

## METHOD[1]

1. I obtained a camera and two types of anaglyph glasses (red-cyan and Color-Code).
2. I tested the relative transmittance of the complementary color filters of the anaglyph glasses (see Table 5):
   a. I shone light from the flashlight at the NXT light sensor through the red and then the cyan filters of a pair of red-cyan glasses and noted the sensor readings.
   b. I repeated for various red-cyan and Color-Code glasses (the latter not shown in the results because red cyan proved to yield better 3D perception).
3. I wrote an image-processing program to generate red-cyan and Color-Code anaglyphs:
   a. I installed Python 2.6.6 and PIL 1.1.7 and patched the image split function.[2]
   b. I set up folders for sources and outputs.
   c. I wrote the program to convert 2D images into 3D anaglyphs, and save in *Sources* folder.
   d. I incorporated in the program the measured factor to boost the intensity of the color layer whose

[1] All the figures referred to in this Section are illustrated in the Appendix.
[2] http://hg.effbot.org/pil-2009-raclette/changeset/fb7ce579f5f9#top

corresponding  filter had a lower reading.

gen metric.py - E:\Python\SciProj\Sources\gen metric.py

*Code snippet to generate Color-Code anaglyph*

```
import Image,ImageChops
def am(im,im2,name,c):
    im.convert("RGB")
    im2.convert("RGB")
    r,g,b=im.split()
    r2,g2,b2=im2.split()
    im2 = im2.convert("L")
    b2=b2.point(lambda i: i * 1.5)
    im3=Image.merge("RGB",(r,g,b2))
    im3.save("..\Outputs\om/"+name)
    print name,"was successfully saved."

def im(x,c=0):
    im=Image.open(x+".jpg")
    rts=[]
    amtl=[1.3,3.8,6.4,8.9,11.4]
    for amt in amtl:
        rts.append([Image.open("R"+x[1:]+"_"+str(amt)+".jpg"),"R"+x[1:]+"_"+str(amt)+".png"]
    for i in range(len(rts)):
        item=rts[i]
        amt=amtl[i]
        if c:
            am(im,item[0], x+"R"+x[1:]+"_"+str(amt)+"mon.jpg",c)
        else:
            am(im,item[0], x+"R"+x[1:]+"_"+str(amt)+".jpg",c)
        amt=amt+2.54
```

Ln: 24 Col: 0

gen metric.py - E:\Python\SciProj\Sources\gen metric.py

*Code snippet to generate red-cyan anaglyph*

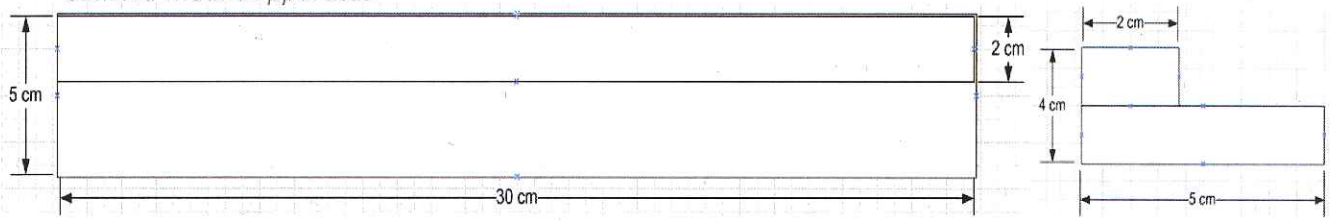
```
import Image,ImageChops
def am(im,im2,name,c):
    im.convert("RGB")
    im2.convert("RGB")
    im=im.point(lambda i: i * 1.2)
    r,g,b=im.split()
    r2,g2,b2=im2.split()
    im2 = im2.convert("L")
    im3=Image.merge("RGB",(r,g2,b2))
    im3.save("..\Outputs\om/"+name)
    print name,"was successfully saved."

def im(x,c=0):
    im=Image.open(x+".jpg")
    rts=[]
    amtl=[1.3,3.8,6.4,8.9,11.4]
    for amt in amtl:
        rts.append([Image.open("R"+x[1:]+"_"+str(amt)+".jpg"),"R"+x[1:]+"_"+str(amt)+".png'
    for i in range(len(rts)):
        item=rts[i]
        amt=amtl[i]
        if c:
            am(im,item[0], x+"R"+x[1:]+"_"+str(amt)+"mon.jpg",c)
        else:
            am(im,item[0], x+"R"+x[1:]+"_"+str(amt)+".jpg",c)
        amt=amt+2.54
```

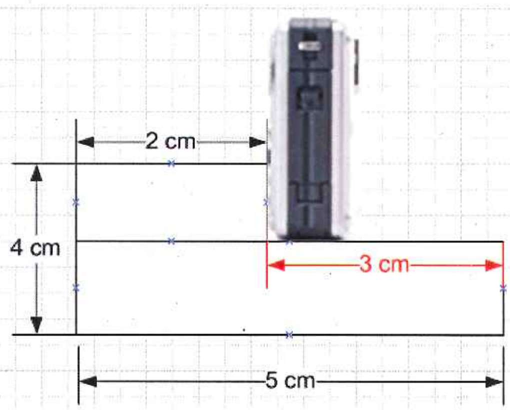
Ln: 24 Col: 0

4. I built a camera-mount apparatus with scale as follows:
    a. I attached a 30 cm × 2 cm × 2 cm piece of wood on top of a 30 cm × 5 cm × 2 cm piece of wood.

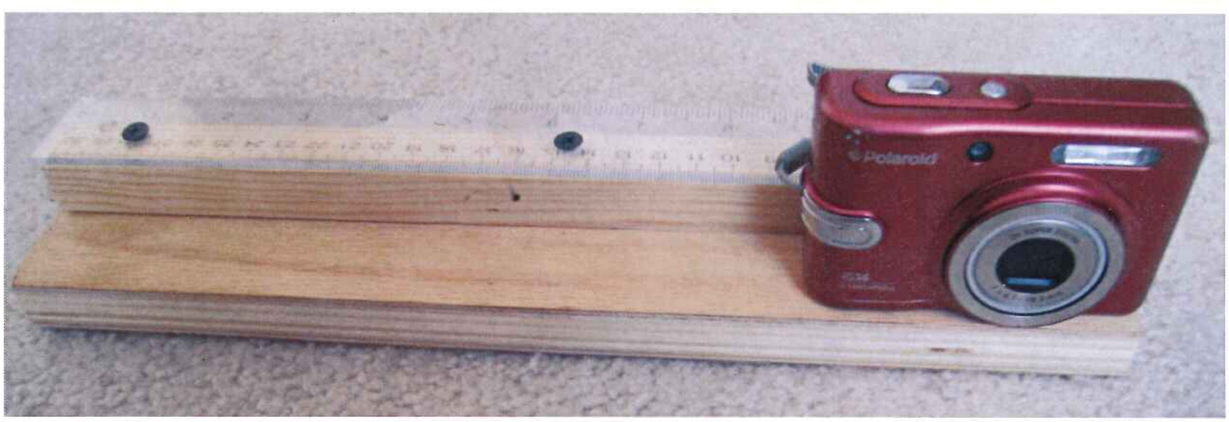
*Camera-mount apparatus*




    b. Then, I attached a standard ruler on its flat side on top of the upper plank of wood.
    c. I made sure the camera can be placed on the 3-cm extension with some camera feature (*e.g.*, left boundary of the LCD screen) at the at the 0-cm mark of the scale.




*Camera placement on camera-mount apparatus*



5. Using the above apparatus, I took one picture (called it the first picture) with the camera at the zero mark.
6. I moved the camera 1.3 cm right, 3.8 cm right, 6.4 cm right, 8.9 cm right, 11.4 cm right, of the original zero position and took five additional pictures (called them second, third, fourth, fifth, and sixth pictures).
7. I uploaded the pictures to the computer and moved them into the *Sources* folder.
8. I named the first picture: L1.jpg, the second R1_1.3.jpg, the third R1_3.8.jpg, the fourth R1_6.4.jpg, and so on.
9. I right clicked the program file and clicked "Edit with IDLE".
10. When IDLE opened, I clicked the window with the program code in it, and pressed "F5".
11. In the shell (other window), I typed: im ("L1"), and pressed "Enter" to generate the anaglyphs.
12. The anaglyphs got generated in the outputs folder.
13. I performed steps 5 through 13 with eight different images (objects or scenes). The images, however, needed different filenames than the first set. The second set's first image was named L2, and the rest of the images in the second set were named identically to the similar ones in the first set, except the R1 in the filename became an R2. The filenames were coined in a way that the left-right separation levels of the different anaglyphs were mentioned in those names

14. I ran a survey on an audience of 20 subjects (with a mix of boys, girls, adult men, adult women); I explained to and asked the subjects to use the ITU BT.500 5-point quality scale to score the images of all sets, and collected the data on pre-printed survey sheets.

15. I picked a representative image of each set (at a particular separation) and asked the subjects to separately score the red-cyan and the Color-Code versions of the image.

16. I copied the data from the survey sheets to the computer and analyzed the results to determine: the ideal separation (results of Step 15) and the difference between the two different types of anaglyph glasses (results of Step 16).

*Example survey sheet*

Separation Testing Using Red-Cyan Glasses

| Separation | Set 1 | Set 2 | Set 3 | Set 4 | Set 5 | Set 6 | Set 7 | Set 8 |
|---|---|---|---|---|---|---|---|---|
| 1.3 cm | 3 | 4 | 5 | 4 | 5 | 5 | 5 | ~~4~~ 5 |
| 3.8 cm | 4 | 4 | 5 | 4 | 5 | 5 | 5 | 5 |
| 6.4 cm | 4 | 4 | 5 | 5 | 4 | 5 | 4 | 5 |
| 8.9 cm | 3 | 4 | 4 | 5 | 4 | 5 | 5 | 5 |
| 11.4 cm | 4 | 5 | 4 | 5 | 4 | 5 | 4 | 5 |

Glasses Comparison

| Rating For: | Set 1 | Set 2 | Set 3 | Set 4 | Set 5 | Set 6 | Set 7 | Set 8 |
|---|---|---|---|---|---|---|---|---|
| Red-Cyan | 4 | 4 | 5 | 5 | 4 | 5 | 4 | 5 |
| Color-Code | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 3 |

# Results

I photographed eight sets of pictures with my digital camera, each set comprising photos at six different camera positions. Using the first position (in each set) as a reference for the other five, five anaglyphs are generated per set for five different separations, and their quality is evaluated using the ITU BT.500 5-point scale using 20 subjects. These # of anaglyphs per set × # of sets × # of subject evaluations = 5 x 8 x 20 = 800 data points are then grouped, averaged, plotted, and interpreted.

Table 2 and the graph in Figure 3 show the average score over all sets of photos for each separation of the red-cyan anaglyphs. Table 3 and the graph in Figure 4 show a comparison of the average scores over all sets of photos for a particular separation (6.4 cm) of the red-cyan and the Color-Code anaglyphs. Figure 5-Figure 12 along with Table 4 show the scoring for individual sets. Table 5 shows the transmittance of various red-cyan glasses. Figure 13 and Figure 14 present two red-cyan anaglyph examples (from two of the sets used in this experiment) for viewing with the glasses enclosed.

**Table 2. Average red-cyan score chart over all sets over all subjects**

| | Average overall score | | | | |
|---|---|---|---|---|---|
| Separation | 1.3 cm | 3.8 cm | 6.4 cm | 8.9 cm | 11.4 cm |
| Score | 3.40 | 3.67 | 3.74 | 3.52 | 3.05 |

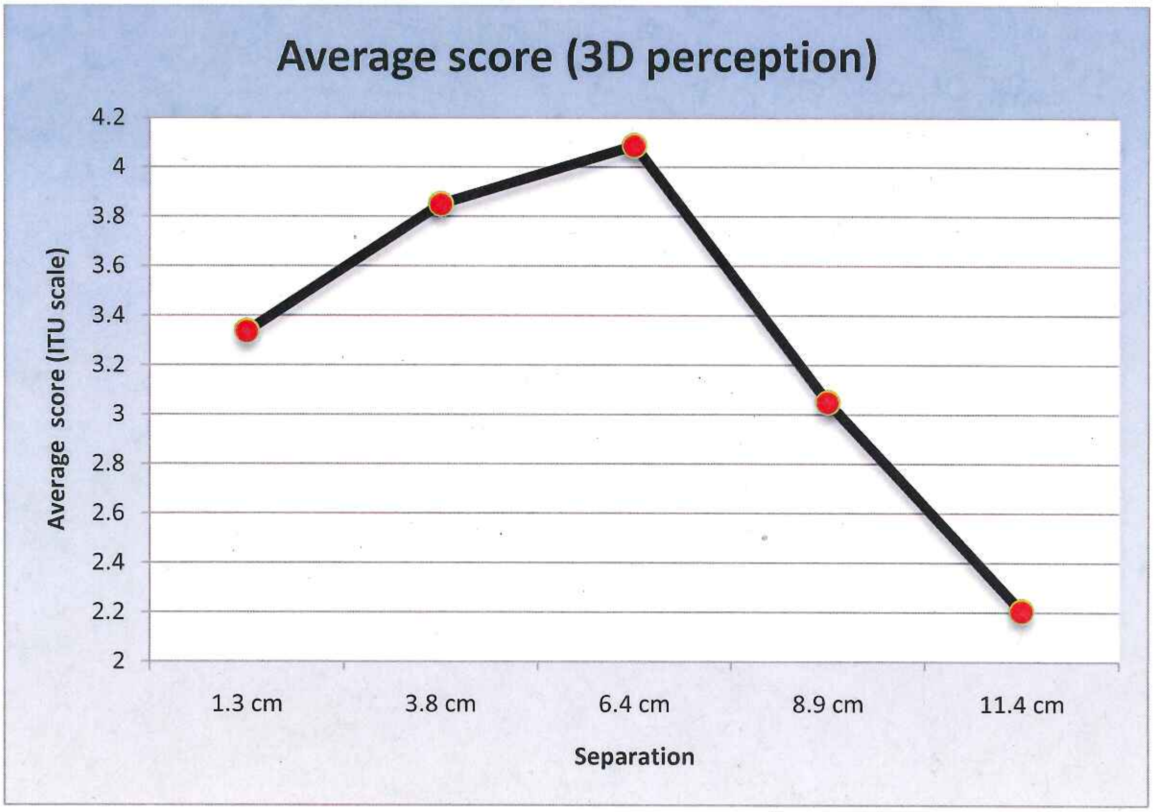



**Figure 3. Average 3D perception score over all sets (from Table 2)[3]**

[3] A line graph is used to emphasize the drop off in score as the separations are farther apart than 6.4 cm.

**Table 3. Color-Code vs. Red-cyan average scores over all sets at a fixed separation of 6.4 cm**

| | Average scores of Color-Code vs. Red Cyan | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Set 1 | Set 2 | Set 3 | Set 4 | Set 5 | Set 6 | Set 7 | Set 8 |
| Color-Code | 2.65 | 3.60 | 3.15 | 3.10 | 2.80 | 3.00 | 2.55 | 2.95 |
| Red-Cyan | 3.50 | 3.75 | 3.45 | 3.45 | 3.20 | 3.10 | 2.75 | 3.45 |

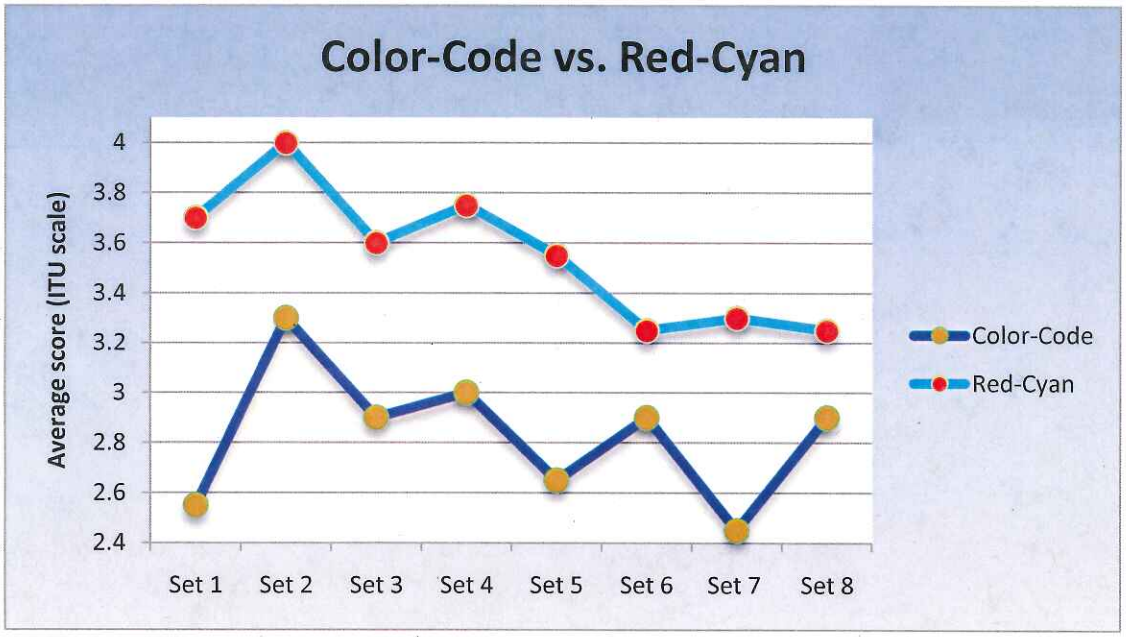



**Figure 4. Color-Code versus Red-cyan average scores over all sets (from Table 3)**

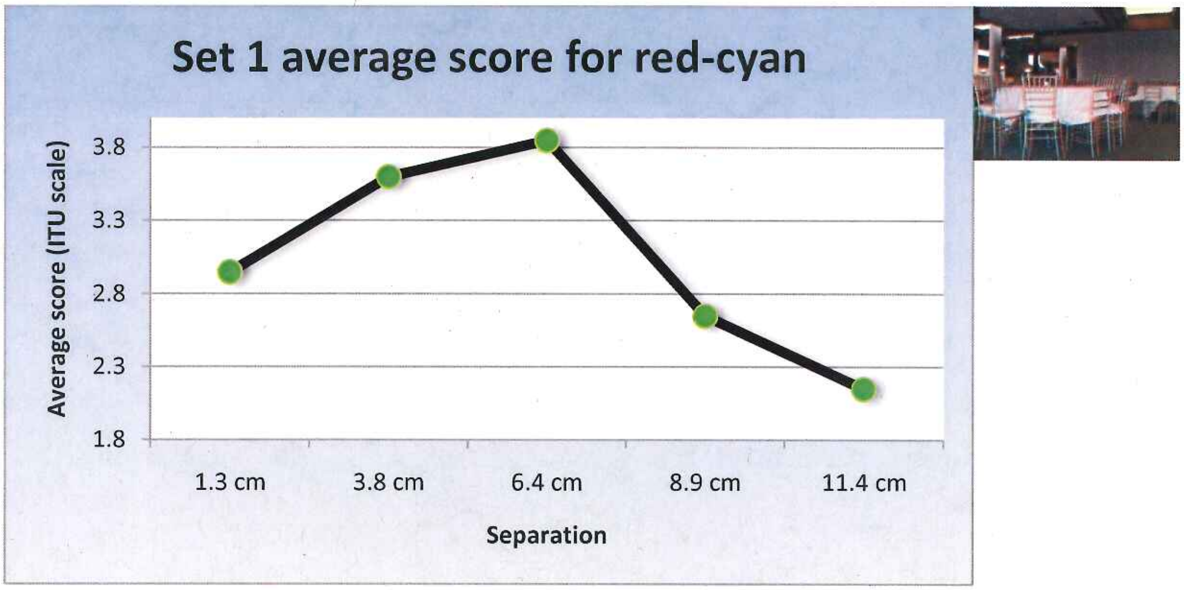



Figure 5. Average score for Set 1 red-cyan data over 20 subjects (from Table 4)

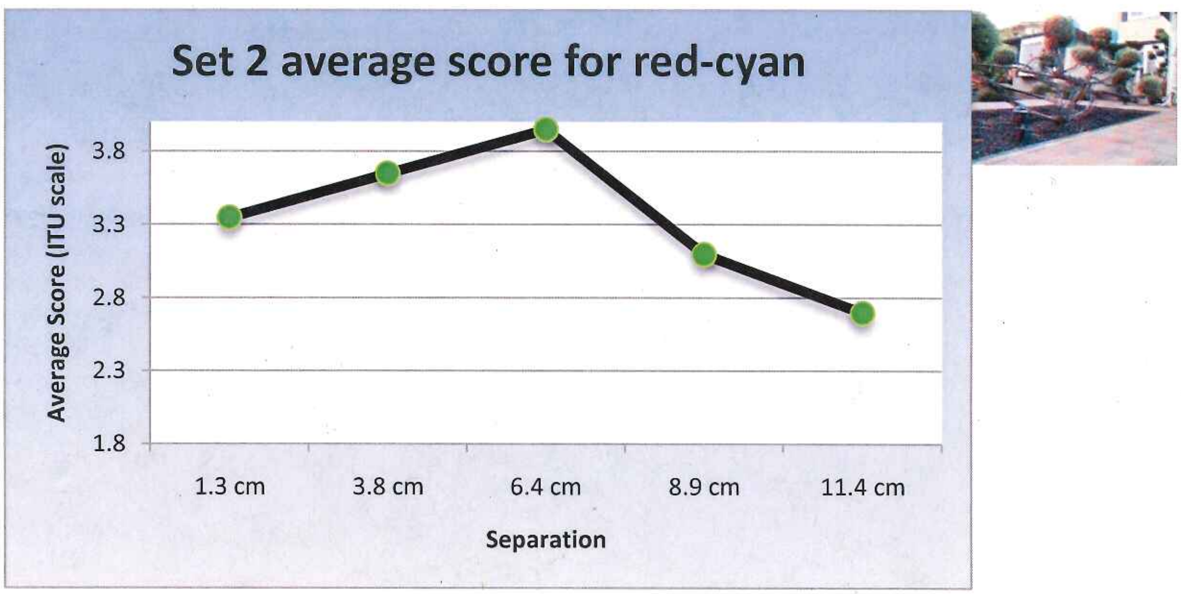



Figure 6. Average score for Set 2 red-cyan data over 20 subjects (from Table 4)

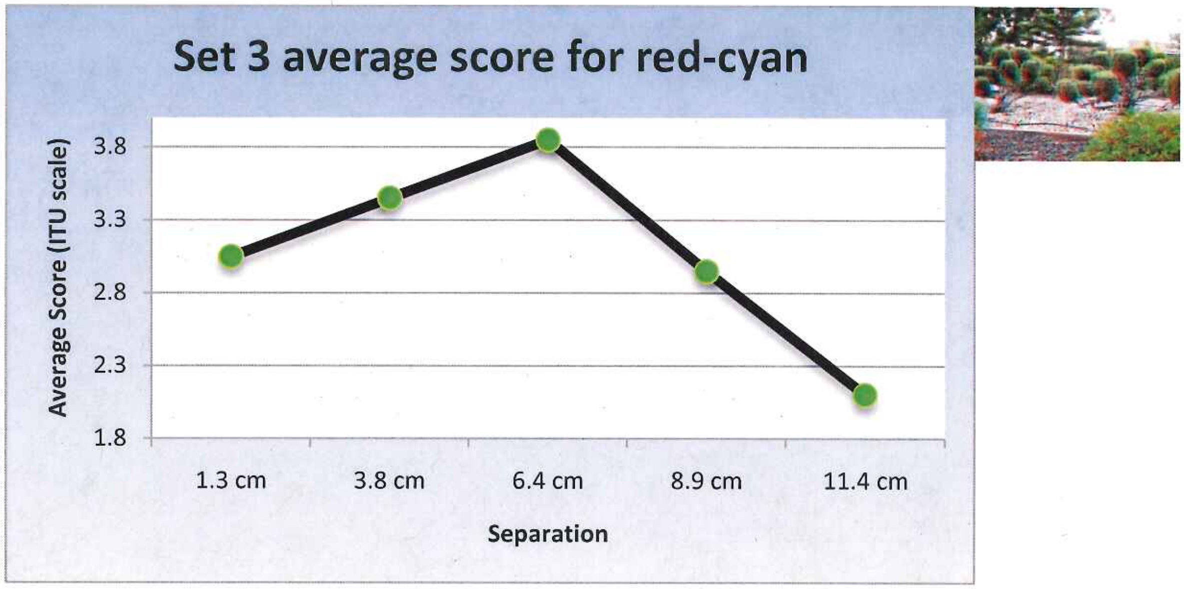



**Figure 7. Average score for Set 3 red-cyan data over 20 subjects (from Table 4)**

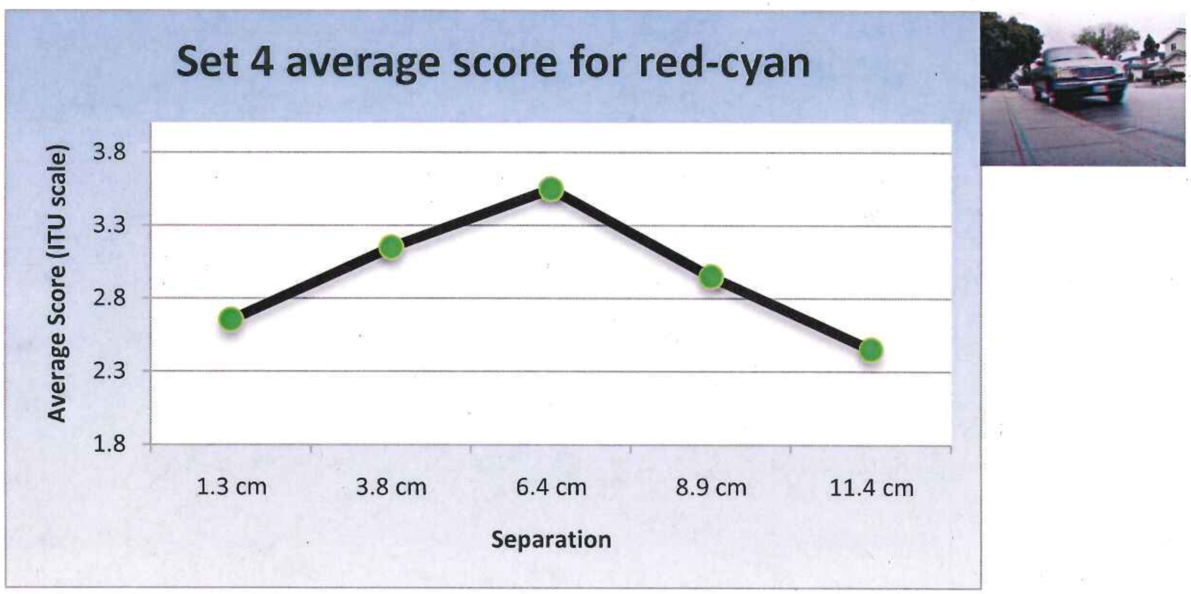



**Figure 8. Average score for Set 4 red-cyan data over 20 subjects (from Table 4)**

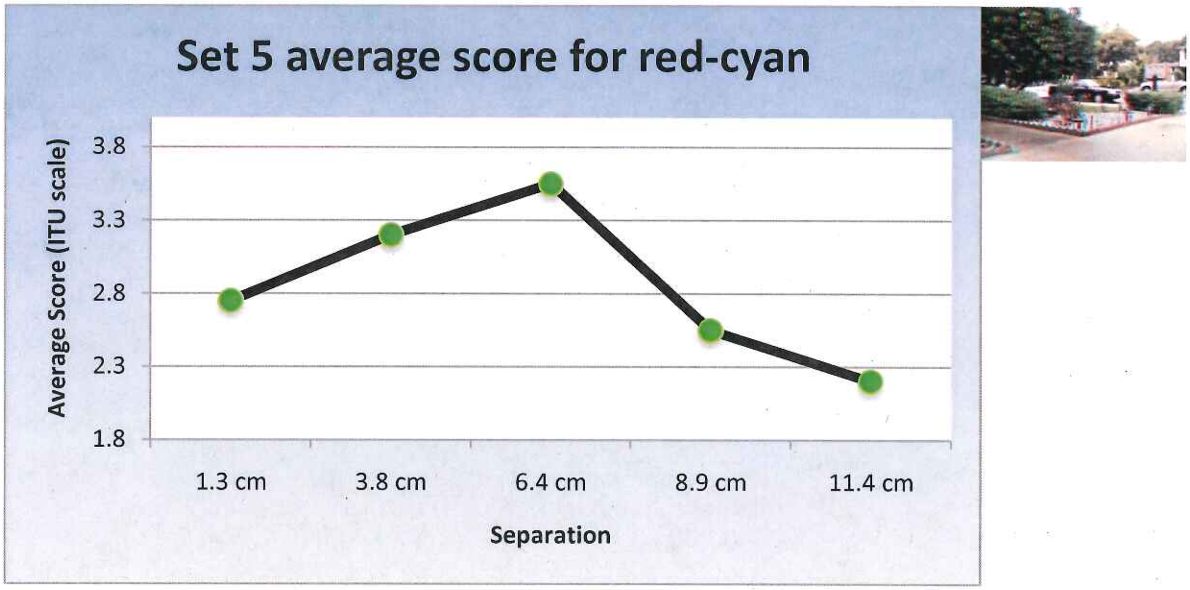



**Figure 9. Average score for Set 5 red-cyan data over 20 subjects (from Table 4)**

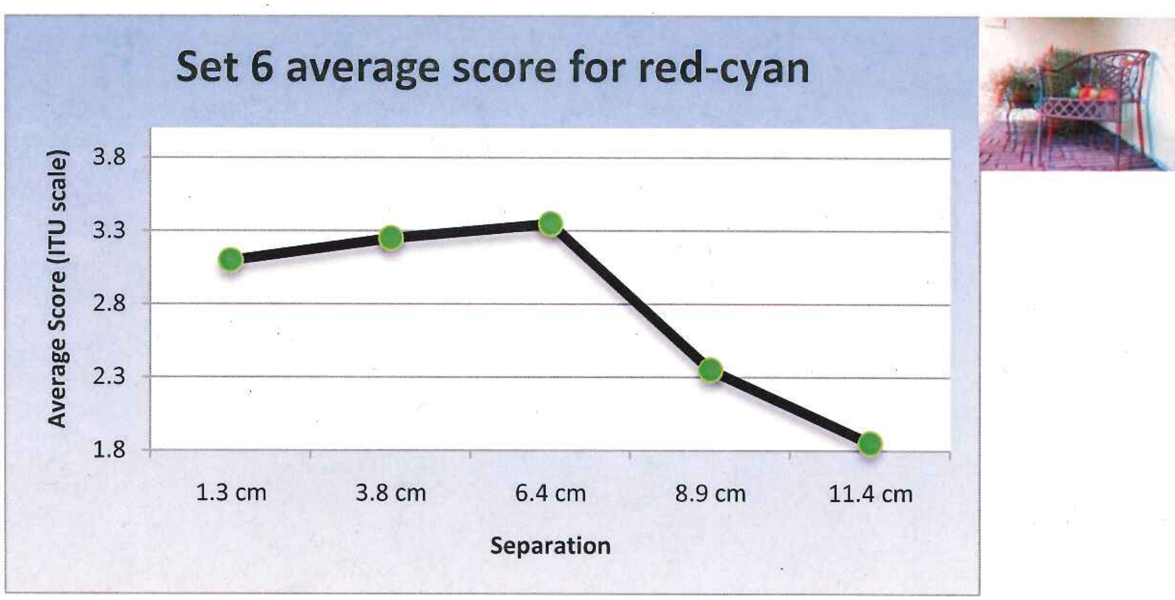



**Figure 10. Average score for Set 6 red-cyan data over 20 subjects (from Table 4)**

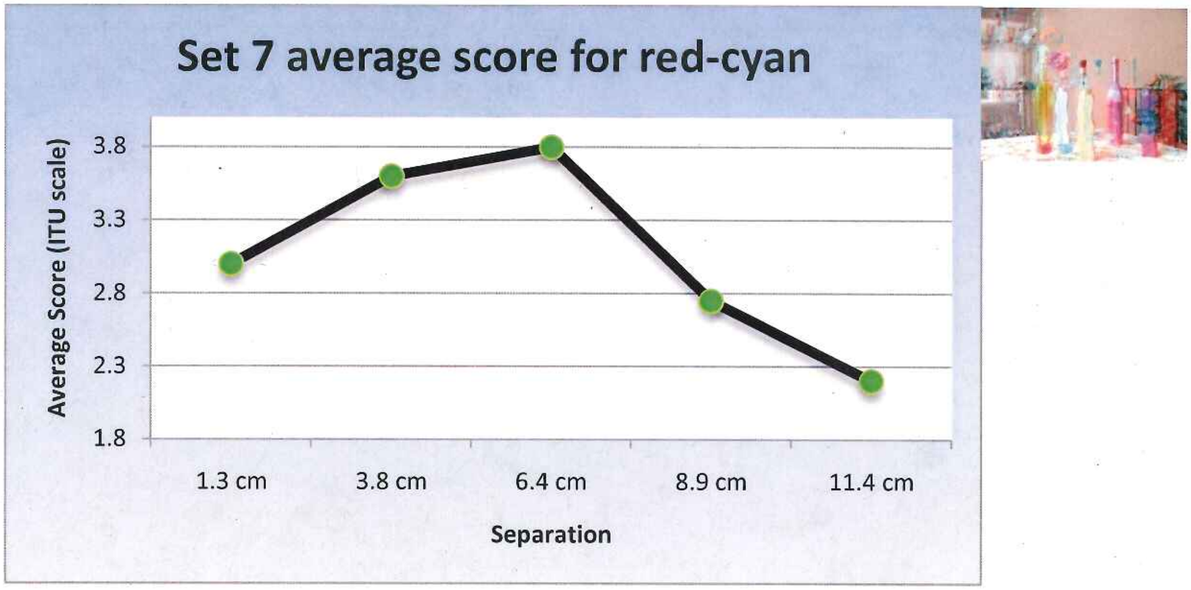



Figure 11. Average score for Set 7 red-cyan data over 20 subjects (from Table 4)

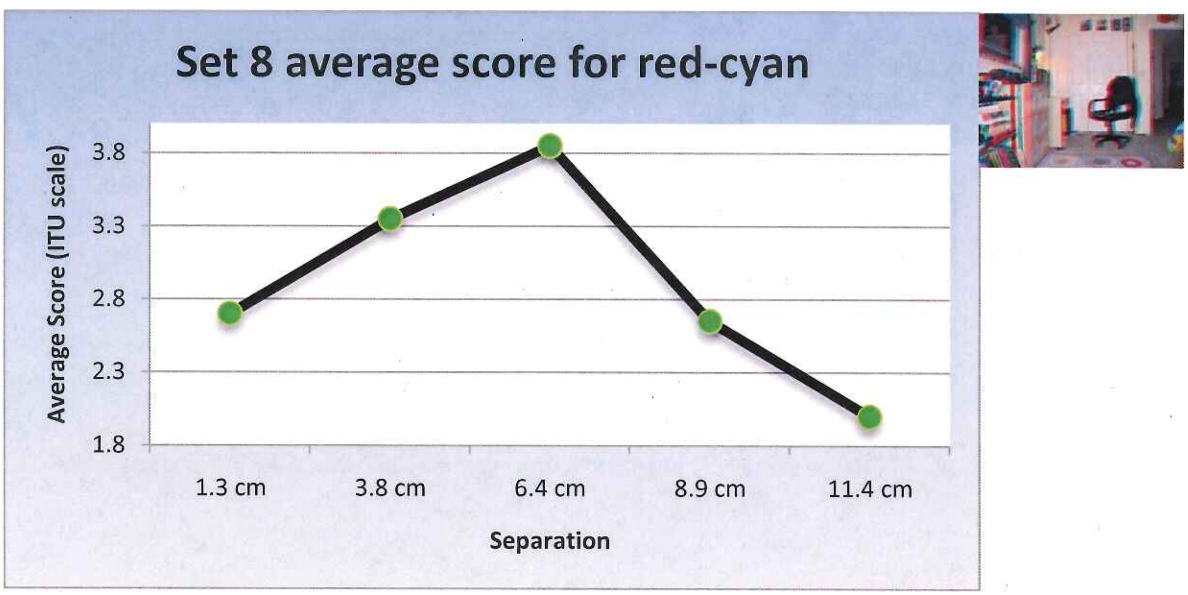
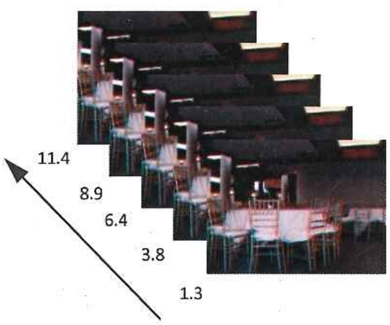



Figure 12. Average score for Set 8 red-cyan data over 20 subjects (from Table 4)

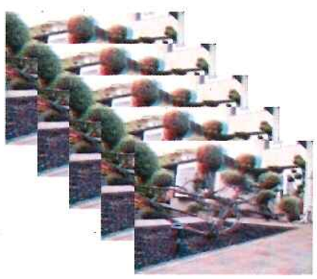
Table 4. Average score for different separations for different sets of pictures over 20 subjects

| | Set averages for various separations for red-cyan glasses | | | | |
|---|---|---|---|---|---|
| | 1.3 cm | 3.8 cm | 6.4 cm | 8.9 cm | 11.4 cm |
| Set 1 | 2.95 | 3.60 | 3.85 | 2.65 | 2.15 |
| Set 2 | 3.35 | 3.65 | 3.95 | 3.10 | 2.70 |
| Set 3 | 3.05 | 3.45 | 3.85 | 2.95 | 2.10 |
| Set 4 | 2.65 | 3.15 | 3.55 | 2.95 | 2.45 |
| Set 5 | 2.75 | 3.20 | 3.55 | 2.55 | 2.20 |
| Set 6 | 3.10 | 3.25 | 3.35 | 2.35 | 1.85 |
| Set 7 | 3.00 | 3.60 | 3.80 | 2.75 | 2.20 |
| Set 8 | 2.70 | 3.35 | 3.85 | 2.65 | 2.00 |
| Avg. | 3.33 | 3.85 | 4.09 | 3.05 | 2.20 |

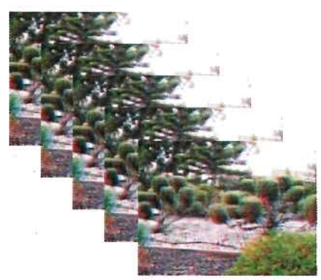
Set 1



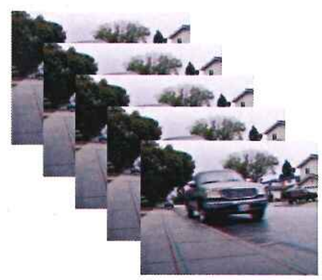

Set 2



Set 3



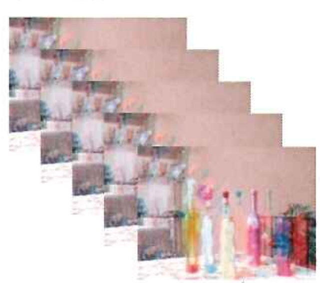
Set 4



Set 5

Set 6

Set 7



Set 8

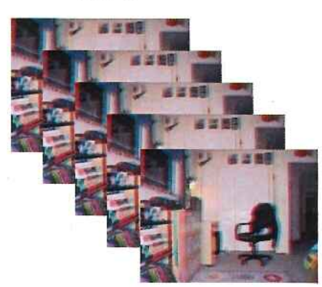


Table 5. Amount of transmittance for various glasses

| Transmittance levels for red-cyan filters | | | |
|---|---|---|---|
| Glasses | NXT sensor reading | | Intensity boost needed for red |
| | Red | Cyan | |
| Minoru 1 | 63 | 75 | 19% |
| Minoru 2 | 75 | 90 | 20% |
| Standard 1 | 62 | 74 | 19% |
| Standard 2 | 65 | 78 | 20% |
| Avg. | 62.5 | 74.75 | 20% |

Figure 13. Example 1: Red-cyan anaglyph (my neighbor's house) ... *view with enclosed glasses*

Figure 14. Example 2: Red-cyan anaglyph (front of our house) ... *view with enclosed glasses*

## Conclusions

The results largely match my hypothesis. However, I learnt several new things as well.

Good quality 3D anaglyphs were produced using a standard digital camera and taking multiple images by shifting the camera horizontally. The camera shift worked well as long as the camera was moved only laterally and the subject stayed still.

The overall 3D perception (score) was better for the red-cyan anaglyphs. However, the Color-Code 3D anaglyphs yielded better reds. The cyan filter on the 3D glasses let in more light than the red filter. So, the pixels of the red image layer needed to be given a 20% intensity boost before the anaglyphs were generated.

The most comfortable overall 3D experience was at a separation of 6.4 cm, approximately the distance between the human eyes. The scores were quite close for the separations of 3.8 cm and 6.8 cm, but rapidly fell off after that. So, humans favor low separation and depth than high ones --- a flatter picture is better than one with so much depth that it is difficult for the eye to resolve.

In future, I would like to figure out how to generate good 3D anaglyphs even when there is relative motion of objects in the scenes captured, or when there is camera motion; this would allow people to take two photos without any stand (using just their camera phones) and generate 3D. It would also be interesting to try to estimate depth and generate anaglyphs from a single image by artificially producing left and right views via image cropping.

## Acknowledgements

I would like to thank my science teacher, Mrs. Morgensen, for allowing me to do this project and for her editorial suggestions to improve the report.

I would like to thank my father for helping me enter my data and proofread my report for errors.

I would like to thank my mother for helping me put together the display board.

# Bibliography


1. Calvert, J. B. Stereopsis. 25 November 2010 <http://mysite.du.edu/~jcalvert/optics/stereops.htm>.
2. Helga Kolb, Eduardo Fernandez, and Ralph Nelson. The Perception of Depth. 28 November 2010 <http://webvision.med.utah.edu/KallDepth.html>.
3. Magic Eye, Inc. What is stereo vision? 1995. 28 November 2010 <http://www.vision3d.com/stereo.html>.
4. Marr, John. Stereopsis. 1982. 28 November 2010 <http://pages.slc.edu/~ebj/sight_mind/stereo/stereo.html>.
5. Oracle Inc. . Binocular Disparity. 28 November 2010 <http://library.thinkquest.org/27066/depth/nldisparity.html>.
6. Panasonic. Introduction . 28 November 2010 <http://fullhd3d.panasonic.eu/en_GB/article/6/1282/introduction-to-3d-perception>.
7. Powell, John M. Stereopsis. 28 November 2010 <http://www.mikesflightdeck.com/stereopsis.htm>.
8. Teittinen, Marko. Depth Cues in the Human Visual System. 25 November 2010

<http://www.hitl.washington.edu/scivw/EVE/III.A.1.c.DepthCues.html>.

9. Wattie, John. How to make anaglyphs in Photoshop. 28 November 2010 <http://nzphoto.tripod.com/sterea/anaglyph_make.html>.


M I N O R U

Red-cyan glasses
enclosed on last
page.