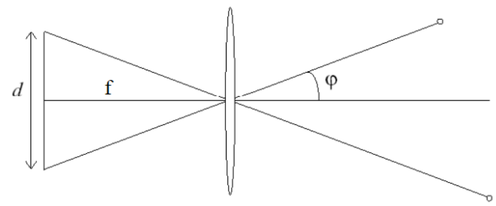
CSC528 Assignment #1. Alex Teboul

Problem 1: CCD to Camera Transformation
Consider a perfect perspective projection camera with focal length 24 mm and a CCD array of size 16 mm × 12 mm, containing 500 × 500 pixels.

Field of View (FOV) is defined as the angle between two points at opposite edges of the image (CCD array), either horizontally or vertically. Thus there are two FOVs, one horizontal and one vertical. The FOV is twice the angle between the optical axis and one edge of the image.

i. Give a general expression for computing horizontal FOV from focal length and image width.



, where: $\varphi = \tan^{-1}(\frac{d}{2f})$

^Vision Lecture

- From this diagram, I assume *d* to be the image width, *f* to be the focal length, and φ to be ½ the FOV.
- $Horizontal\ FOV = 2*\varphi$
- $Horizontal\ FOV = 2*(\ tan^{-1}\left(\frac{d}{2f}\right)\ )$

$$Horizontal\ FOV = 2*tan^{-1}(\frac{image\ width}{2*focal\ length})$$

ii. Compute the horizontal FOV and vertical FOV of the given camera.

a. $Horizontal\ FOV$

i. $$Horizontal\ FOV = 2*tan^{-1}(\frac{image\ width}{2*focal\ length})$$

ii. $$Horizontal\ FOV = 2*tan^{-1}(\frac{16mm}{2*(24mm)})$$

iii. $Horizontal\ FOV = 36.86^\circ$

b. $Vertical\ FOV$

i. $$Vertical\ FOV = 2*tan^{-1}(\frac{image\ width}{2*focal\ length})$$

ii. $$Vertical\ FOV = 2*tan^{-1}(\frac{12mm}{2*(24mm)})$$

iii. $Vertical\ FOV = 28.08^\circ$

iii. Comment on how FOV affects resolution in an image.

a. Resolution in a general sense refers to the number of pixels in an image or other proxies for the amount of visual information and detail that can be resolved in an image. Field of view is the angle between two points at opposite edges of the image. By shrinking the field of view, more visual information from a particular section of an image can reach a camera's CCD array. This in turn increases the resolution of that section of the image. In lecture, this concept was discussed in terms of shrinking field of view to zoom in to increase resolution or increase field of view when zooming out to decrease resolution.

Problem 2: Exercise 2.2 from Szelinski book (2D transform editor)
Feel free to use any existing code/libraries you wish, in whatever language you like.

Ex 2.2: 2D transform editor Write a program that lets you interactively create a set of rectangles and then modify their "pose" (2Dtransform). You should implement the following steps:
1. Open an empty window ("canvas").

2. Shift drag (rubber-band) to create a new rectangle.

3. Select the deformation mode (motion model): translation, rigid, similarity, affine, or perspective.

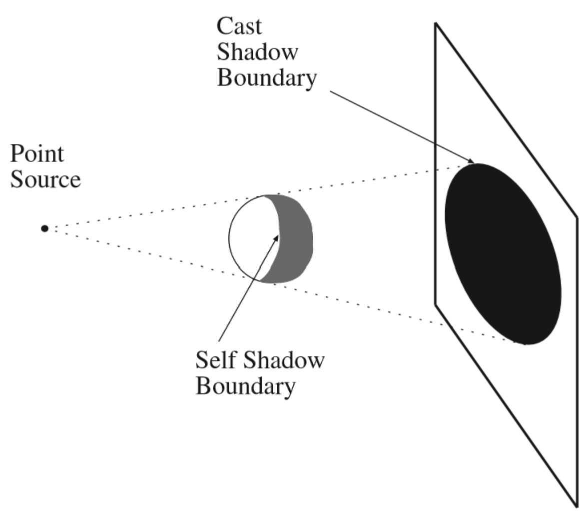
4. Drag any corner of the outline to change its transformation.
This exercise should be built on a set of pixel coordinate and transformation classes, either implemented by yourself or from a software library. Persistence of the created representation (save and load) should also be supported (for each rectangle, save its transformation).

Problem 3: Exercises from Forsyth book

i. 2.8.1 #1

What shapes can the shadow of a sphere take, if it is cast on a plane, and the source is a point source?

a. Spheres project onto planes as circles and ellipses when the light comes from a point source. Changing the angle of the point source relative to the plane or vice versa can change this shape. Perfectly center with the sphere and perpendicular to the plane will yield a circle.



Forsyth book Figure 2.6 showing a point source shown on a sphere and the circle it casts. If the plane were angled differently, the cast shadow could be an ellipse. Same effect with changing the point source angle/direction while keeping the sphere and plane fixed.

i. 2.8.1 #9

9. If one looks across a large bay in the daytime, it is often hard to distinguish the mountains on the opposite side; near sunset, they are clearly visible. This phenomenon has to do with scattering of light by air — a large volume of air is actually a source. Explain what is happening. We have modelled air as a vacuum, and asserted that no energy is lost along a straight line in a vacuum. Use your explanation to give an estimate of the kind of scales over which that model is acceptable.

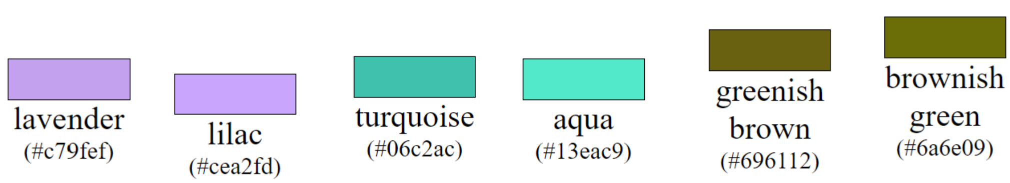
a. In daytime, more light enters the area of the bay and so more light is scattered versus when the sun is setting. As a result, the sky is lighter (and bluer) during the day and there is less contrast between the mountains and the sky during the day. An increased contrast between the mountains and sky at sunset would make them easier to see. If the mountains are snowcapped, they're also going to be reflecting more light, compounding this effect. If the mountains are across bay, there's also an increased likelihood of humidity in the air decreasing the maximum line of sight. Colder temperatures also improve maximum line of sight. General temperature and humidity changes between midday and sunset/sundown could also be influencing this phenomenon.

For the average person standing up at roughly 5'7" to 6', the horizon is a little over 3 miles away across the flatness of the bay. The mountains are higher above the horizon though, so they can be seen quite a long way away. Models of the air that assume it's a vacuum tend to work fine for most non-technical applications depending on the weather. For example, from a tall building you could still take photos of mountains that are up to 100miles away on a clear day. It's not really fair to estimate the scales over which the model is acceptable because it depends so greatly on the environment. For example, saying it works at distances up to a mile is clearly not true if there's fog, but is fine it's a clear day.

Clearly, over shorter distances, the model of air as a vacuum is more acceptable than over long distances. Factors that influence its applicability are: humidity level, particulate matter in the air, air temperature, ambient light intensity or time of day (ex. overcast vs clear skies and daytime vs sunset). With low humidity, low particulate matter in the air, cold air temperature, clear skies, and nearing sunset the air as a vacuum model is most acceptable. With these conditions, it could apply at distances in the miles range, up to the horizon. With the opposite conditions, like high humidity, lots of particulate matter, high air temp, lots of ambient light and scatter, etc. the air as a vacuum might only hold up in the feet to inches ranges – depending on their levels.

ii. 3.7 #1 (Then read **https://blog.xkcd.com/2010/05/03/color-survey-results/**)

Sit down with a friend and a packet of coloured papers, and compare the colour names that you use. You will need a large packet of papers — one can very often get collections of coloured swatches for paint, or for the Pantone colour system very cheaply. The best names to try are basic colour names — the terms red, pink, orange, yellow, green, blue, purple, brown, white, gray and black, which (with a small number of other terms) have remarkable canonical properties that apply widely across different languages [?; ?; ?]. You will find it surprisingly easy to disagree on which colours should be called blue and which green, for example.



These were some of the colors that my friend an I disagreed on. I called the lavendar and lilac purple, while my friend called them pink. The turquoise and aqua I called blue, while my friend called them green. And the greenish brown and brownish green I called both green, while my friend called them both brown. Most of the other ones we said the same colors though. Used the list from that survey: **https://xkcd.com/color/rgb/**

iii. Chapter 6 exercises, #5 (this has little to do with CV, but is very useful to understand)

> Now consider a single coin that we flip many times, and where each flip is independent of the other. We set up an event structure that does not reflect the order in which the flips occur. For example, for two flips, we would have:
>
> $$\{\emptyset, D, \mathtt{hh}, \mathtt{tt}, \{\mathtt{ht}, \mathtt{th}\}, \{\mathtt{hh}, \mathtt{tt}\}, \{\mathtt{hh}, \mathtt{ht}, \mathtt{th}\}, \{\mathtt{tt}, \mathtt{ht}, \mathtt{tt}\}\}$$
>
> We assume that $P(\mathtt{hh}) = p^2$; a simple computation using the idea of independence yields that $P(\{\mathtt{ht}, \mathtt{th}\}) = 2p(1-p)$ and $P(\mathtt{tt}) = (1-p)^2$. We can generalise this result, to obtain
>
> $$P(k \text{ heads and } n-k \text{ tails in } n \text{ flips}) = \begin{pmatrix} k \\ n \end{pmatrix} p^k (1-p)^{n-k}$$

Example 6.10: *The probability of various frequencies in repeated coin flips*

A careless study of example 10 often results in quite muddled reasoning, of the following form: I have bet on heads successfully ten times, therefore I should bet on tails next. Explain why this muddled reasoning — which has its own name, the gambler's fallacy in some circles, anti-chance in others — is muddled.

a. Each coin flip is independent of past coin flips, so getting 10 coin flips in a row to be heads does not *theoretically* raise the probability of the next flip being tails. It's muddled reasoning to assume that the next flip would yield a tails or that it's at a higher probability of being tails. That said, flipping a coin 10 times in a row and getting heads every time is a low probability of 1/1024. I would actually bet tails if I were in the gambler's position too.

Problem 4: Computer Vision Concepts
Briefly explain the following concepts

i. Perspective projection
   a. Perspective projection refers to the linear projection of 3 Dimensional objects or the visual space onto a 2D image plane. So a ball projected onto a plane or an image of railway lines, in which the lines appear to disappear in the distance. Note that distant objects appear smaller. A vanishing point like this is an important concept for perspective projection.

ii. Vanishing point
   a. A vanishing point is the point at which parallel lines seen in perspective projection appear to converge in the distance. Technically, multiple vanishing points can be used in the projection of objects and scenes onto a plane.

iii. Stereopsis
   a. Stereopsis generally refers to using multiple visual axes that can converge on a particular point. Humans take advantage of this with our two eyes to help us perceive depth. Two slightly different images arrive at each eye in our case, and these images from different perspectives can then be processed to help in the perception of depth.

iv. Optical flow
    a. Optical flow refers to the change in the image projected onto the retina or a camera sensor over time when there is motion. As an object moves or the perspective changes due to the movement of the retina/camera sensor, there is said to be optical flow. This is related to why we can perceive motion in the presentation of still frames presented in sequence (like in a movie).

v. Parallax
    a. Parallax refers to perceived change in position of an object when viewed from different perspectives. If you move side to side, close objects will appear to move quickly in the visual space, while distant objects move more slowly. In terms of parallax, it can be said that nearby objects display a larger parallax than far away objects.

Put all your work into a single file: output images; text responses; and program code. Submit using the dropbox in D2L.