Lecture 1: Introduction to multiple view geometry

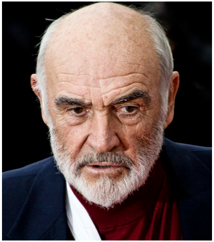
24 January, 2020

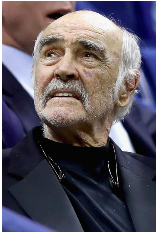
Problem 1. Image reconstruction

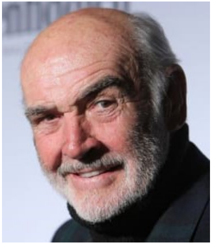
Suppose we are given a number of photographs of a single 3D object. Can we reconstruct the object?

Clearly for a solid object we need more than one photograph (otherwise we can't see the back of the object). Therefore we need multiple views of the same object.

We also need some way of matching the points in one photograph to points in the other photographs.

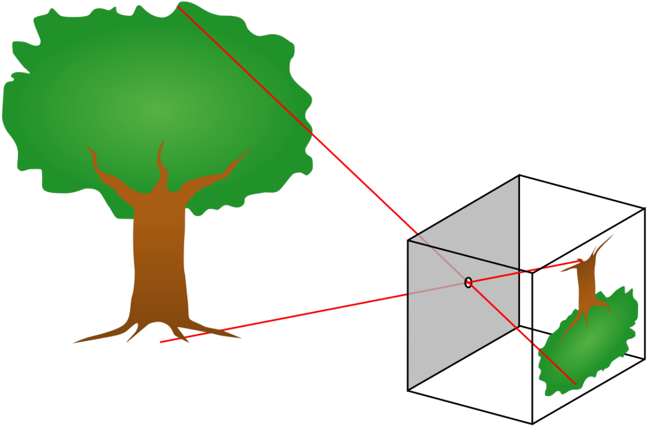






Example: Picture of a 3D object

The following picture shows a simple model of a camera. The 3-dimensional tree is projected through the pinhole onto the screen at the back of the camera.

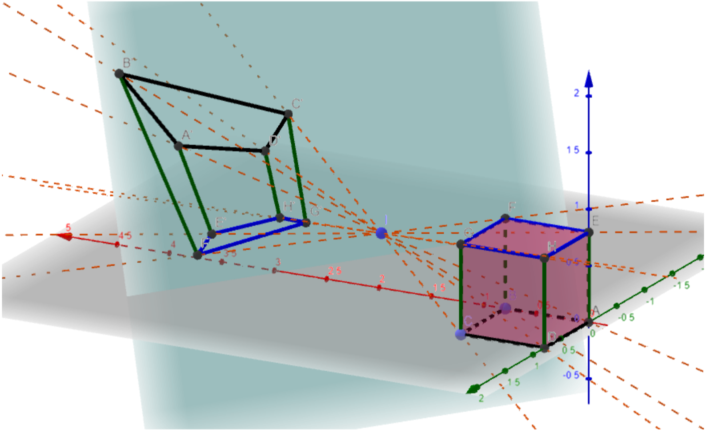


http://commons.wikimedia.org/wiki/Image:Pinhole-camera.png.

Example: Picture of a 3D object

We need to see multiple images of the tree in order to see it as a 3-dimensional object.

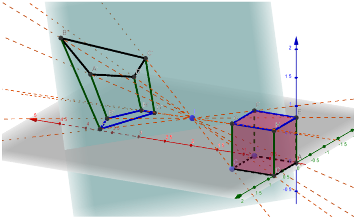
Here is another example of a simple object (a cube) projected through a pinhole onto a screen.

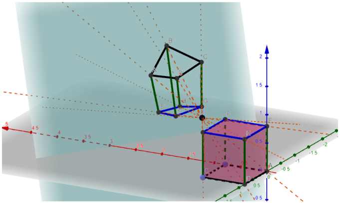


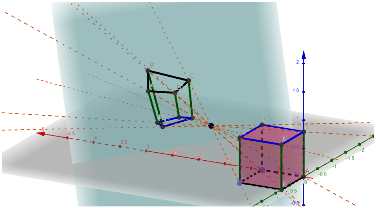
Example: Picture of a 3D object

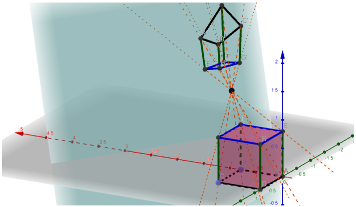
To see the 3-dimensionality of the cube, we need multiple images.

In this case the 2D screen is the same, but the pinhole (or "point of perspective") moves.





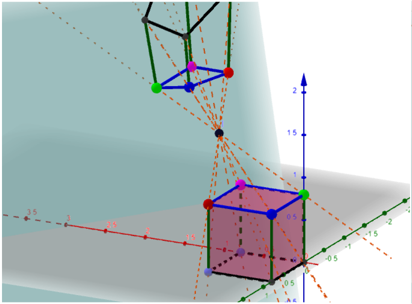


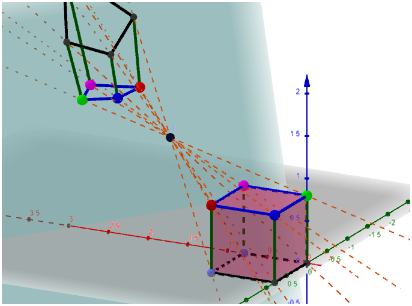


Example: Point matching between different photographs of the same object

To reconstruct the 3-dimensional cube, we also need to know which points on the cube match up with each point in the projection.

In the example below, we have matched up four of the points (red, blue, green, pink) in the first photograph with four points in the second photograph.





More examples: Trick photographs

You may have already seen some examples of how the brain can be tricked by a single 2D picture of a 3D object.

The first video shows a picture of what looks to be an "impossible" staircase, where the stairs all go upwards, but they return to the original starting point. As the camera angle moves, we can see that this is an illusion caused by the projection of the 3D image onto the 2D screen. Changing camera angle (moving the point of perspective) helps us to see the true nature of the 3D object. https://www.youtube.com/watch?v=erQ4Mqv9l4A.

Another example of this is in the video: "Ascending against gravity"
https://www.youtube.com/watch?v=xVyYcAI90jw.

A more basic example is the "Penrose triangle", in which three blocks of wood form the sides of what looks to be an impossible shape. The link below takes you to a video.
https://www.youtube.com/watch?v=gcw1IIGSGMM.

Some motivating questions (cont.)

Problem 2. **Image recognition**

Suppose that we are given data (e.g. 2D photographs) of many 3D objects, and then we are given a 2D photograph of one of these objects. How can we determine which object is in the photograph?

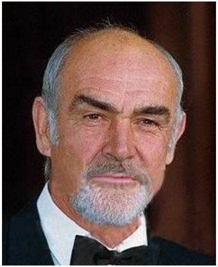
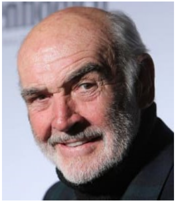
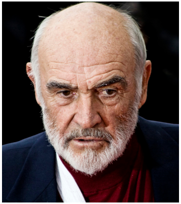
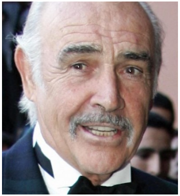
Example. (Facial recognition)

Suppose a computer has stored data on the facial features of a large number of people. Given a photograph of one of these people, how do we match that photograph to one of the people in the system?

A different kind of image recognition problem

Suppose that we are given detailed data on one 3D object (e.g. many 2D photographs), and then a 2D photograph of an unknown object. How do we tell whether it is the object in our database?

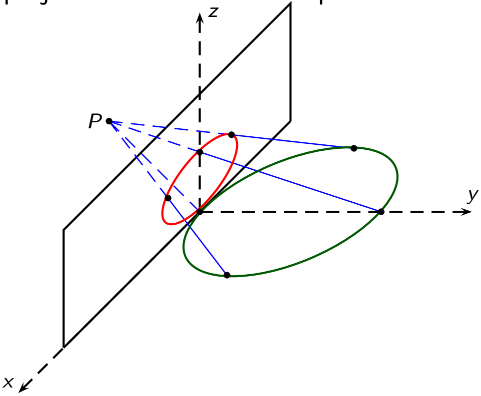
Example. (Facial recognition again)
Suppose your phone has stored data about your facial features (for example, it may have lots of 2D pictures of your face). How does it tell whether the person holding the phone is you or someone else?

In particular, different camera positions may make the same object look different.

Example. In our previous pictures of the cube, each projection onto the 2D plane looks different.

Example. In the picture below, the green circle in the *xy* plane is projected onto the red ellipse in the *xz* plane.



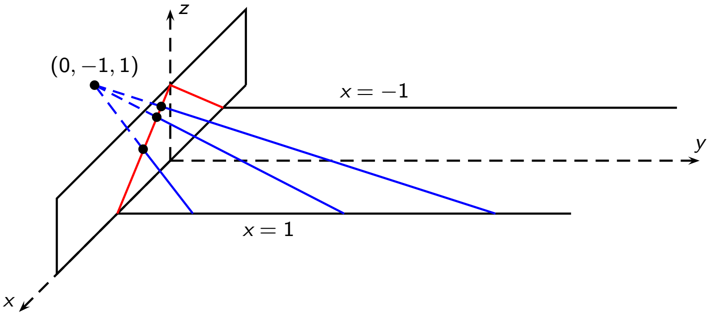
In general, a circle may project onto an ellipse, a parabola or a hyperbola. In order to recognise that the original object is a circle, we need to understand the geometry underlying these projections.

Example: Geometric features may change after projecting

In the following picture, the sides of the airport runway are parallel. Since the picture is a projection of the airport runway onto a 2D screen, then the parallel lines appear to intersect on the horizon.

Example: Geometric features may change after projecting

The following diagram shows how this works. The parallel lines on the xy plane are projected onto the two intersecting red lines on the xz plane.



The human eye/brain is very good at recognising the same object viewed from different viewpoints.

By understanding the geometry underlying this problem, we aim to develop some methods to analyse this using mathematics.

Goal: Use geometry and linear algebra to approach these problems

Of course, building a software engine to reconstruct or recognise 3D objects from 2D photographs is complicated, and currently the subject of much intense research.

Our goal in these lectures is much simpler:

1. The first aim is to describe the geometry of projections from a 3D object to a 2D plane (for example taking a photograph of a 3D object, or projecting an object onto a screen).

2. The second aim is to use linear algebra to describe this geometry, and then to see how this can be used to approach the problems of image reconstruction or image recognition.

Projecting a 3D image onto a 2D plane

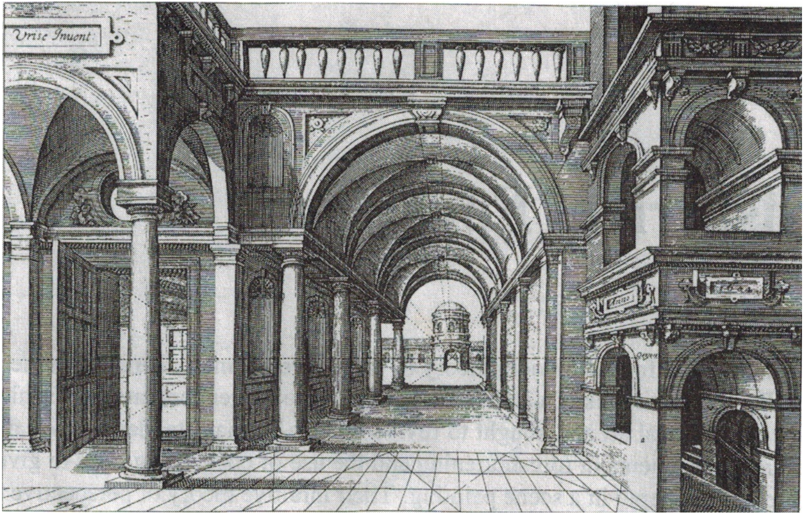
Simple example. Suppose you are standing in a museum and looking at a painting on the wall. Is it an accurately drawn picture?

Equivalently. Is there a place we can stand so that the picture looks realistic? (Does the picture on the wall look the same as a window onto a 3D scene?)

The example on the right is a picture by Albrecht Durer, who was one of the first artists to draw pictures in accurate perspective.

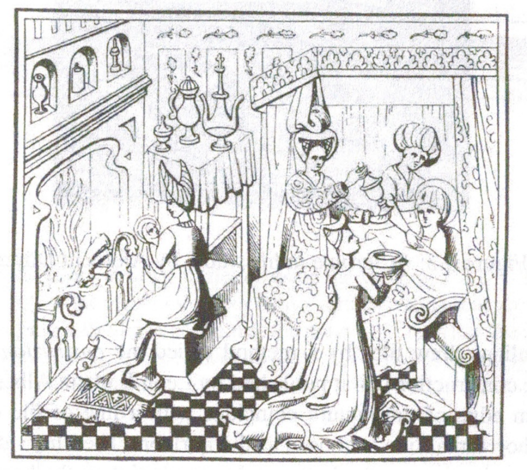
Projecting a 3D image onto a 2D plane

Here is another example of a picture drawn in proper perspective.
Note how the parallel lines meet at a point on the horizon.



Here is an example of a picture which is not drawn in proper perspective.

You can immediately see from the tiled floor that the picture does not look correct.



Drawing in perspective

Before we try to understand how to reconstruct or recognise 3D objects, we first need to understand some of the properties of a picture drawn in perspective, and then understand how that picture can change when we move the point of perspective (equivalently, when we move the camera).

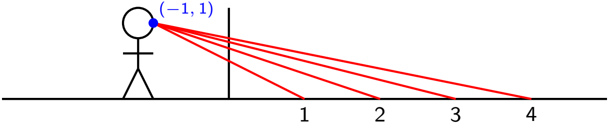
Understanding these properties and the geometry underlying them will help us build a mathematical model of a camera, which we can then use to study the reconstruction and recognition problems.

Problem 0. How to draw an accurate perspective picture?
We will first need to see some examples.

Basic example of projecting onto a 1D line

Example. (Simplified version of viewing a painting on a wall) Suppose you want to draw a picture of the points $1, 2, 3, \ldots, n$ from the x-axis. Once you have finished, you would like to hang your drawing on a wall (the y-axis) and give the viewer (who is positioned at the point $(-1, 1)$) the illusion that they are seeing the x-axis.

Therefore, the points $1, 2, 3, \ldots, n$ (on the x-axis) should be positioned at the points shown below on the y-axis.



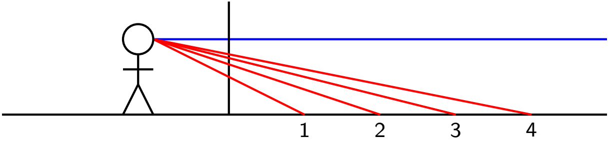
Any problem of this type can be thought of as a projection of one line onto another through a point of perspective.

In this case we are projecting the x-axis onto the y-axis through the point $(-1, 1)$.

Basic example of projecting onto a 1D line (cont.)

The function $y = \frac{x}{x+1}$ is an example of a projection of the x-axis onto the y-axis.

Question. What about $y = 1$? Which point on the x-axis projects to $y = 1$?



We can think of the point $y = 1$ as a point on the horizon. The red line above in the picture above does not intersect the x-axis. Instead we say that it corresponds to a point at infinity, or a point on the horizon.

The point at infinity on the x-axis projects to the point $y = 1$.

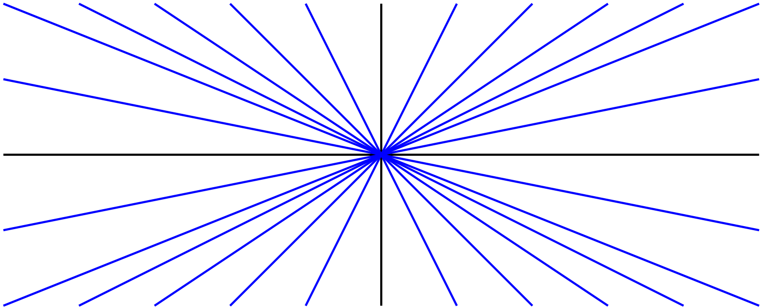
The projective line

Instead of thinking of this example as a projection from the x-axis onto the y-axis, we need to add a point at infinity to both lines.

Below we will define a projective line which includes this point at infinity. Later we will define this projection algebraically.

In the coordinates of $\mathbb{R}^2$, let $(0, 0)$ be our point of perspective.

Definition. The projective line, denoted $\mathbb{R}P^1$, is the set of lines through the origin in $\mathbb{R}^2$.



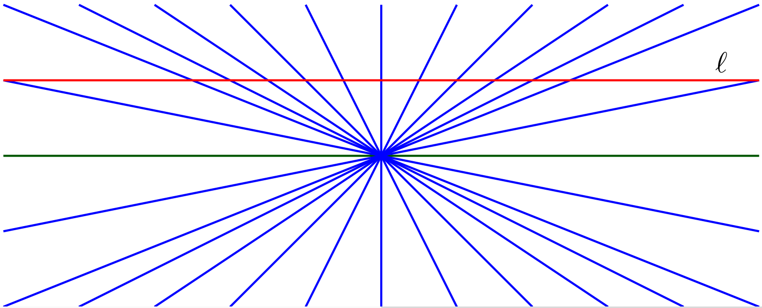
The projective line

Each line through the origin in $\mathbb{R}^2$ corresponds to a point in our projective line.

There are different ways of representing the projective line. Given any line ℓ that doesn't pass through the origin (the red line in the diagram below), each point on $\ell \cup \{\infty\}$ corresponds to a line through the origin.

The point at infinity corresponds to the line parallel to ℓ.

Therefore we can identify $\mathbb{R}P^1$ with $\ell \cup \{\infty\}$.

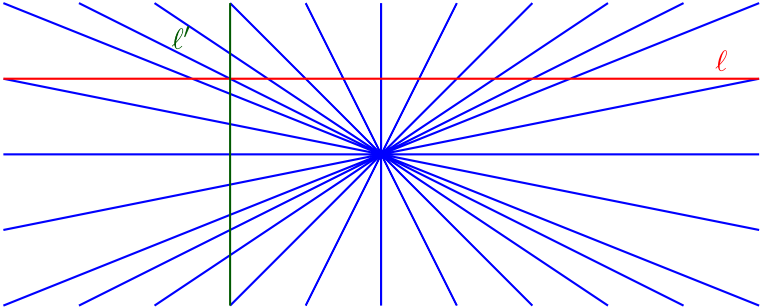


The projective line

If we choose a different line ℓ' which doesn't pass through the origin, then we get a different identification $\mathbb{R}P^1 \cong \ell' \cup \{\infty\}$.

Changing from $\ell \cup \{\infty\}$ to $\ell' \cup \{\infty\}$ is called a projective transformation. This transformation is exactly the projection that we defined earlier.

We will see now how to analyse these transformations algebraically.

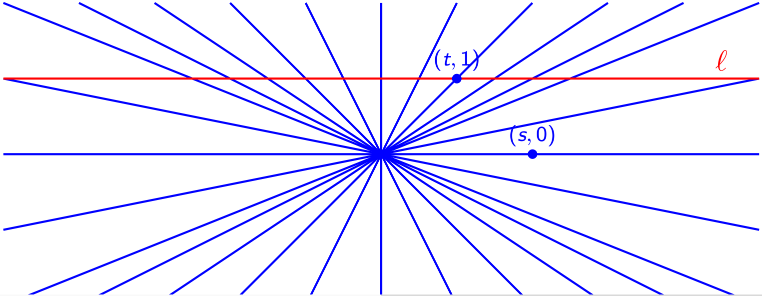


The projective line

Consider the line ℓ in the diagram below. This is the horizontal line through $y = 1$. In coordinates, we can write

$$\ell = \{(t, 1) \mid t \in \mathbb{R}\}.$$

We saw before that the "point at infinity" on the line ℓ corresponds to the horizontal line through the origin. Apart from the origin (the point of perspective) any point on this line can be written as $(s, 0)$ for any $0 \neq s \in \mathbb{R}$. (We will see soon that the choice of s does not matter.)

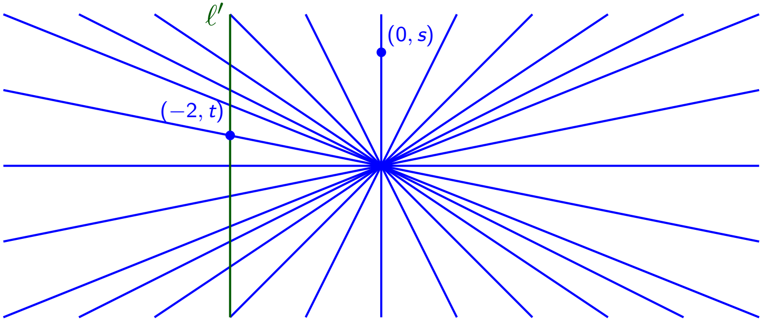


The projective line (cont.)

Now consider another line ℓ' in the diagram below. This is the vertical line through $x = -2$. In coordinates, we can write

$$\ell = \{(-2, t) \mid t \in \mathbb{R}\}.$$

Now the "point at infinity" on the line ℓ' corresponds to the vertical line through the origin (the line through the origin parallel to ℓ'). Apart from the origin (the point of perspective) any point on this line can be written as $(0, s)$ for any $0 \neq s \in \mathbb{R}$.



The projective line (cont.)

How do we relate the points on these two lines?

A point on ℓ is related to a point on ℓ' if and only if they lie on the same line through the origin.

Equivalently, a point (x_1, y_1) on ℓ is related to a point (x_2, y_2) on ℓ' if and only if

$$\frac{x_1}{y_1} = \frac{x_2}{y_2}.$$

(Then both points will lie on the same line $x = \lambda y$.)

In the above equation we allow for $y_1 = 0 = y_2$, in which case both sides are infinity.

A better way to write this relationship between (x_1, y_1) and (x_2, y_2) is to use the following equivalence relation.

$(x_1, y_1) \sim (x_2, y_2)$ if and only if $(x_1, y_1) = \lambda(x_2, y_2)$ for some $\lambda \in \mathbb{R}$.

Definition. The projective line is

$$\mathbb{R}P^1 = \mathbb{R}^2 \setminus \{(0,0)\} / \sim$$

where $(x_1, y_1) \sim (x_2, y_2)$ if and only if $(x_1, y_1) = \lambda(x_2, y_2)$.

Each equivalence class in the projective line corresponds to a line through the origin in $\mathbb{R}^2$.

A convenient way to write an equivalence class is to use homogeneous coordinates, where the line through the origin and the point (x, y) is written as $[x : y]$. These coordinates satisfy the relationship

$$[x_1 : y_1] = [x_2 : y_2] \text{ if and only if } (x_1, y_1) = \lambda(x_2, y_2) \text{ for some } \lambda \in \mathbb{R}.$$

This means that $[x_1 : y_1]$ and $[x_2 : y_2]$ represent the same equivalence class in $\mathbb{R}P^1$ (hence the use of "$=$" instead of "$\sim$" above).

Next time

We will define the projective plane $\mathbb{R}P^2$ as the set of lines through the origin in $\mathbb{R}^3$.

We will then use this to define a model of a camera using matrices and linear algebra.