Lecture 1: Introduction to multiple view geometry

17 January, 2020

Some motivating questions

Problem 1. Image reconstruction

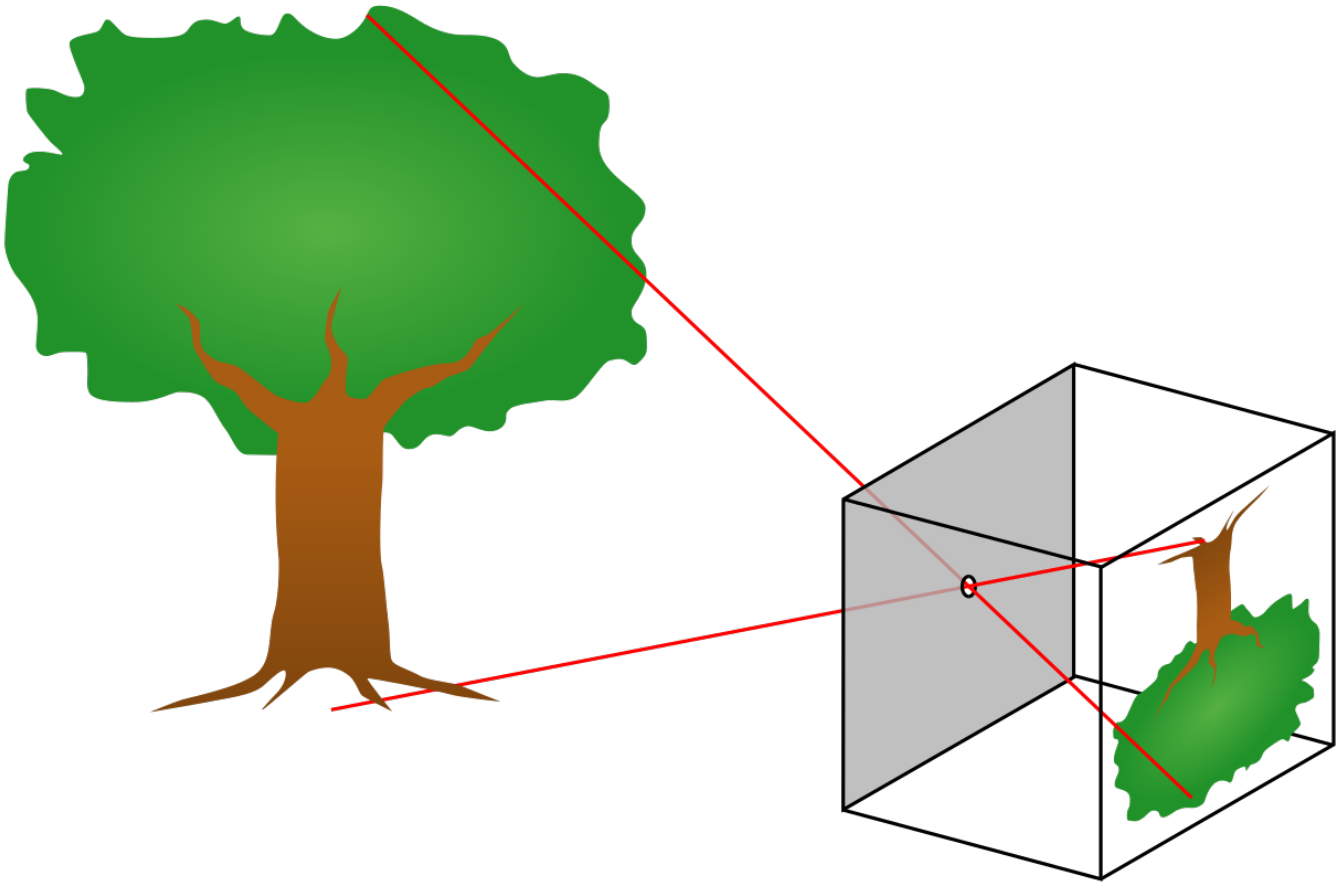
Suppose we are given a number of photographs of a single 3D object. Can we reconstruct the object?

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.

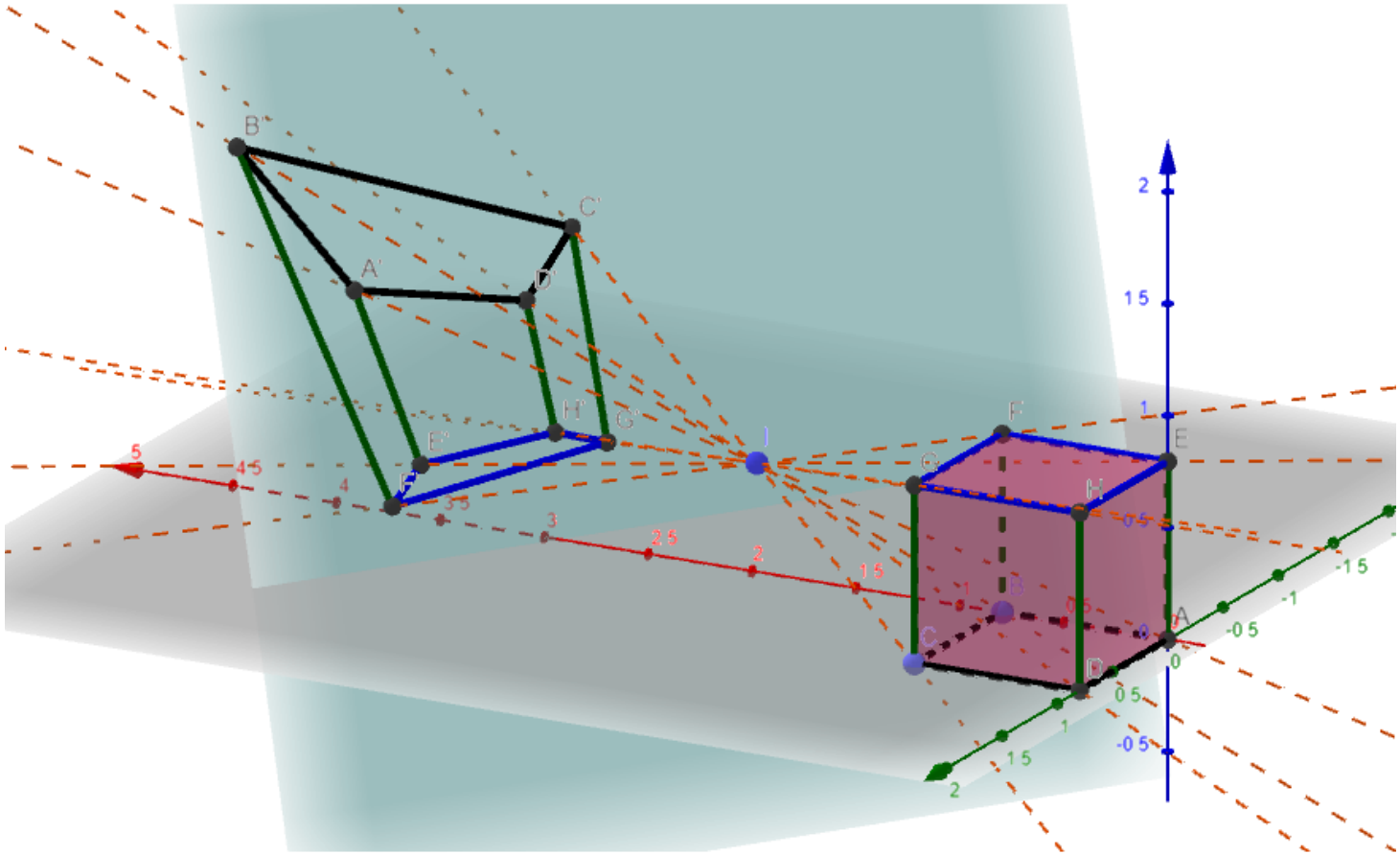


http://commons.wikimedia.org/wiki/Image:Pinhole-camera.png.

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.

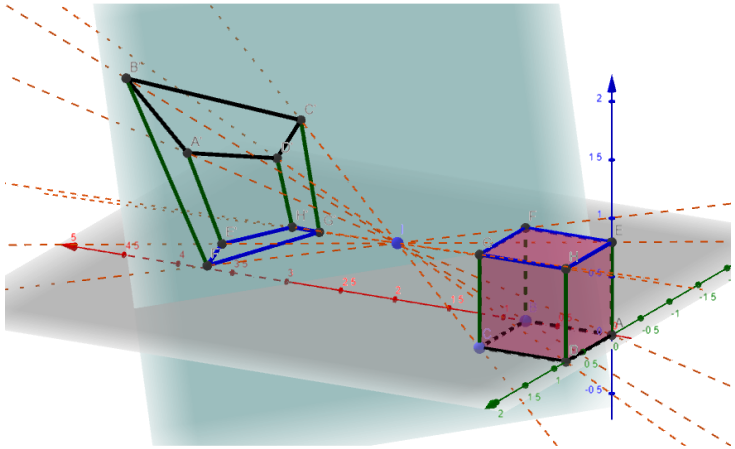
Here is another example of a simple object (a cube) projected through a pinhole onto a screen.

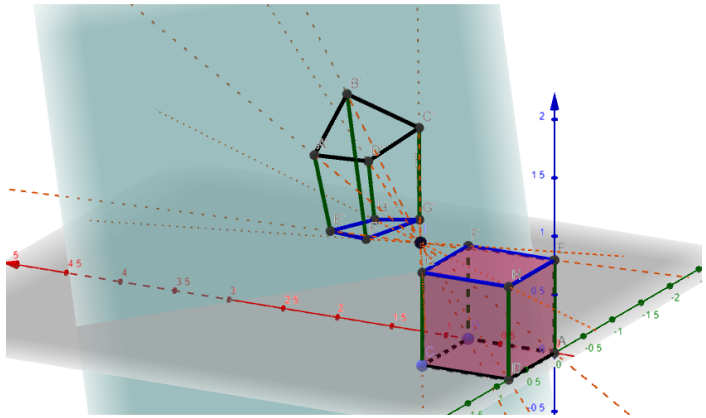


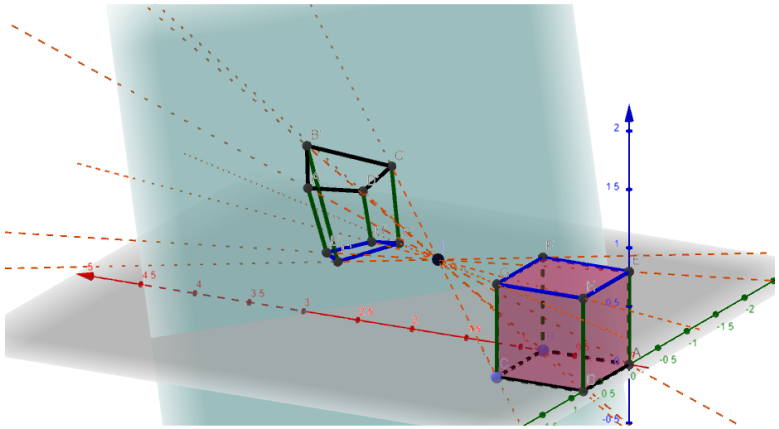
Example: Picture of a 3D object

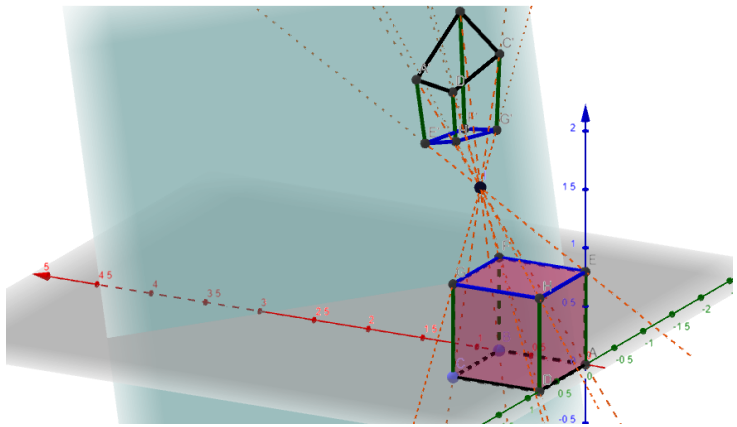
To see the 3-dimensionality of the cube, we need multiple images.

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





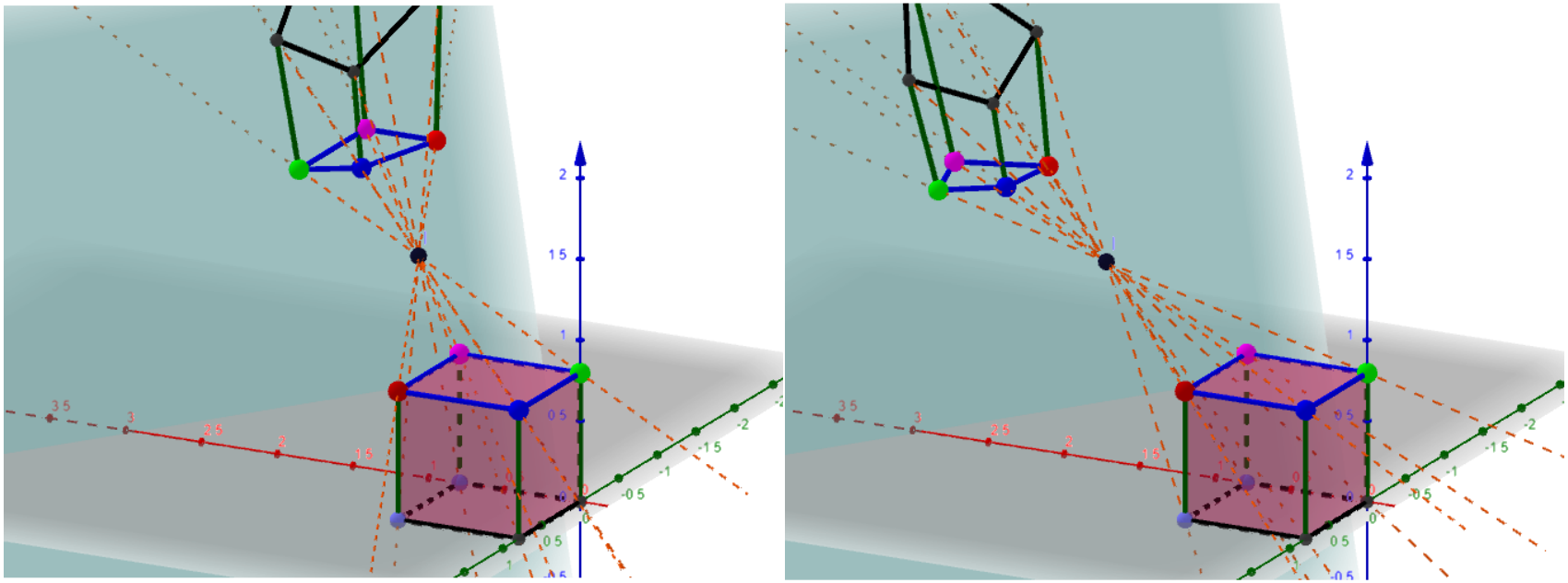




Example: Point matching between different photographs of the same object

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.



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)

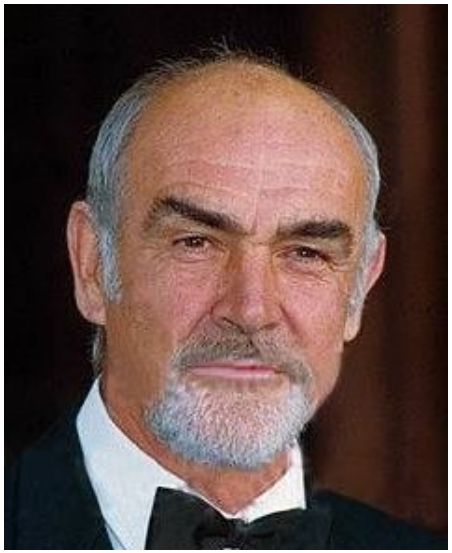
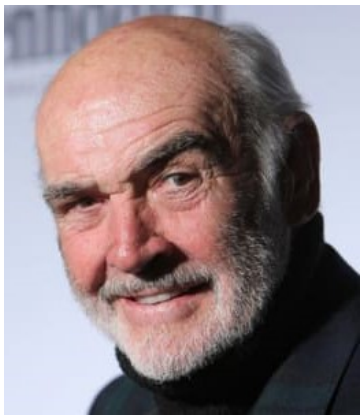
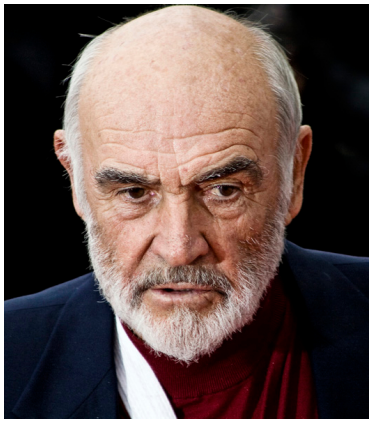
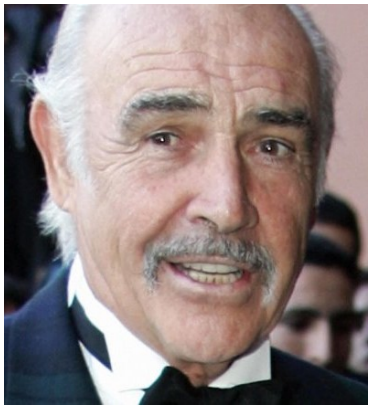
Suppose a computer has stored data on the facial features of a large number of people. Given a photograph of one of these people, how do we match that photograph to one of the people in the system?

Some motivating questions (cont.)

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?

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.

3. In particular, we will see how multiple views of an object can be used to give a sense of the 3-dimensionality of the object.