3D reconstruction from 2D images

Promoters: Jan Lemeire (jan.lemeire@vub.be) & Jan De Beule (Jan.De.Beule@vub.be)

This thesis consists of a mathematical challenge and is related to real-world applications.

Reconstruction of the 3D scene from a single 2D image is an ill-posed problem: information is lost during the perspective transformation. Nonetheless humans are able to perceive depth in images:

A created Illusion

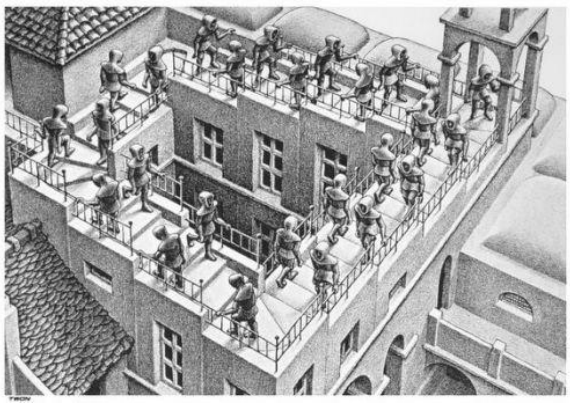
A painting

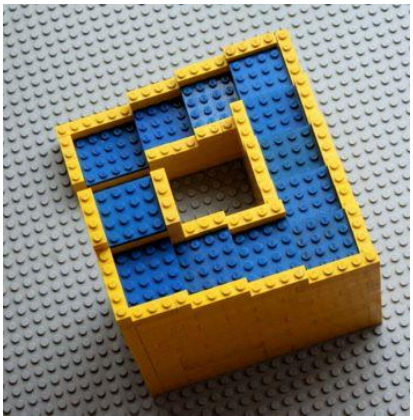
One can make assumptions to tackle the problem. For instance, by assuming that our world is a Manhattan world: the most important lines are along one of the 3 axes X, Y or Z. This is what most research is doing.

The simplicity criterion

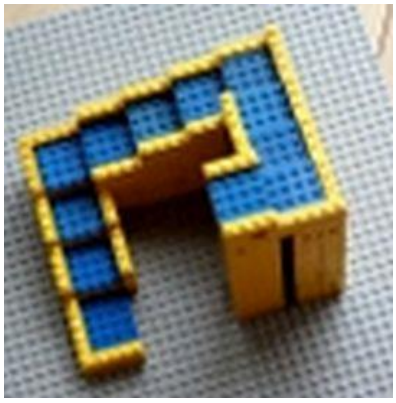
An alternative approach is the application of Occam's Razor: favor simpler models. One searches for the simplest 3D scene that can explain the 2D image. In the left image the human eye assumes that all tiles of the floor are of equal size and shape. However, when simplicity assumptions lead to an inconsistency, the assumption must be revoked, and one should look for a more complex explanation. In the left figure: the tiles are painted on the floor in such a way as to create a 3D illusion.

Consider Escher's staircase:





After some consideration we realize that in the left picture the artist is deliberately deceiving us. While the right picture is misleading us by the angle at which the picture is taken.

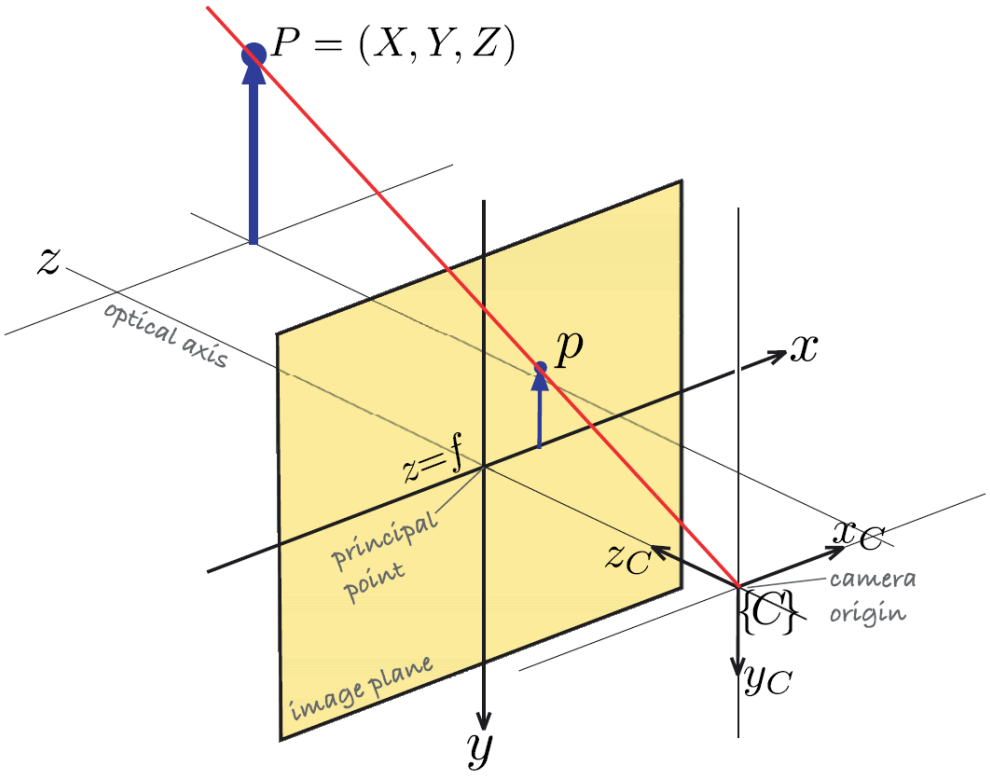


At first sight we assumed that all adjacent blocks were connected, which is a very reasonable assumption. Only by deliberately choosing a *very specific* viewing angle this assumption turns out to be false. For randomly chosen angles (such as in the picture on the left), the assumption is correct, and we correctly infer the 3D structure.

This approach has recently been employed in causal inference (distinguish causes from effects based on pure observational data), another ill-posed problem.

From 3D to 2D

The perspective transform is based on the following projection:



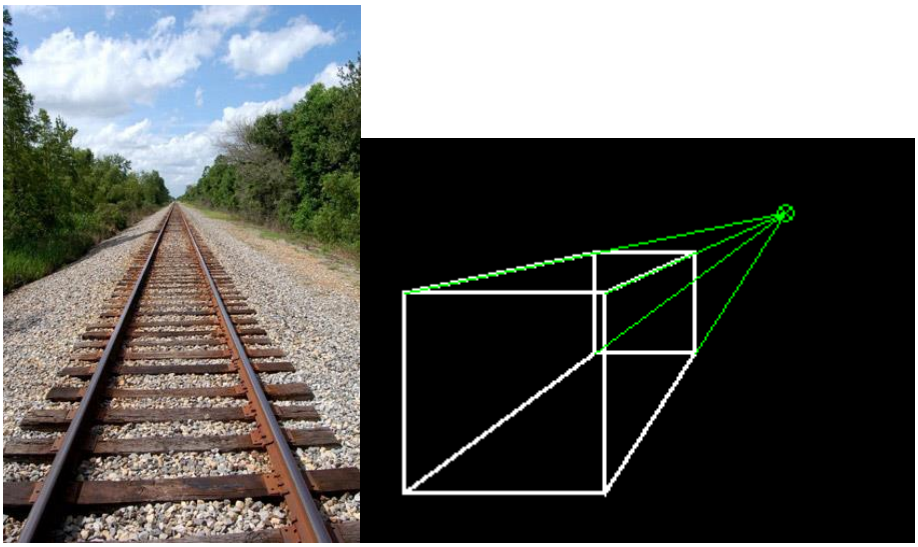
Which leads to the following basic equations:

$$x = f\frac{X}{Z},\ y = f\frac{Y}{Z}$$

All points of the red line (called a 'ray') are projected on the same point of the image. The depth information is lost. This can mathematically be represented by ***homogeneous coordinates***.

Vanishing Points

A basic property from the perspective transform is that parallel lines will cross at infinity or in the same point, called the *vanishing point*. Based on this, the cube structure on the right can be recognized.



Thesis work

At the robotics lab of the department of Industrial Sciences we have developed our own 3D engine in Java. We also have a first implementation of a basic reconstruction algorithm based on vanishing points.

The following should be added to the *mathematical framework* and the *reconstruction algorithm*:

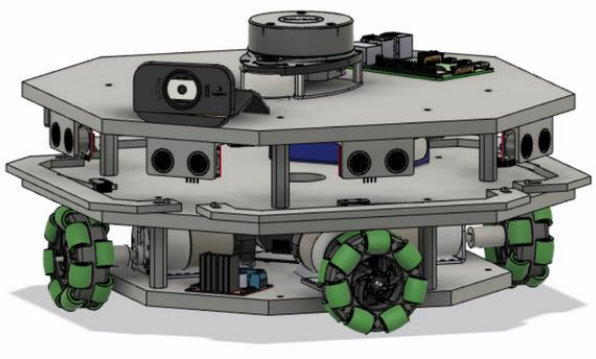
- Use of homogeneous coordinates
- Circular shapes
- Elaboration of the simplicity criterion (e.g. define complexity of structures)

The thesis will start with a literature study. Programming skills are an asset but not required.

Applications

At the robotics lab we focus on 2 applications:

- **Indoor localization:** indoor environments are very 'Manhattan' so ideal for 3D reconstruction from single images. We are building a prototype robot which should be able to navigate indoor.



- **Lego brick recognition**: the bricks have regular shapes which makes them suitable for the proposed approach. We have a sorting machine and a brick grabber.