Visual Pathways to the Brain

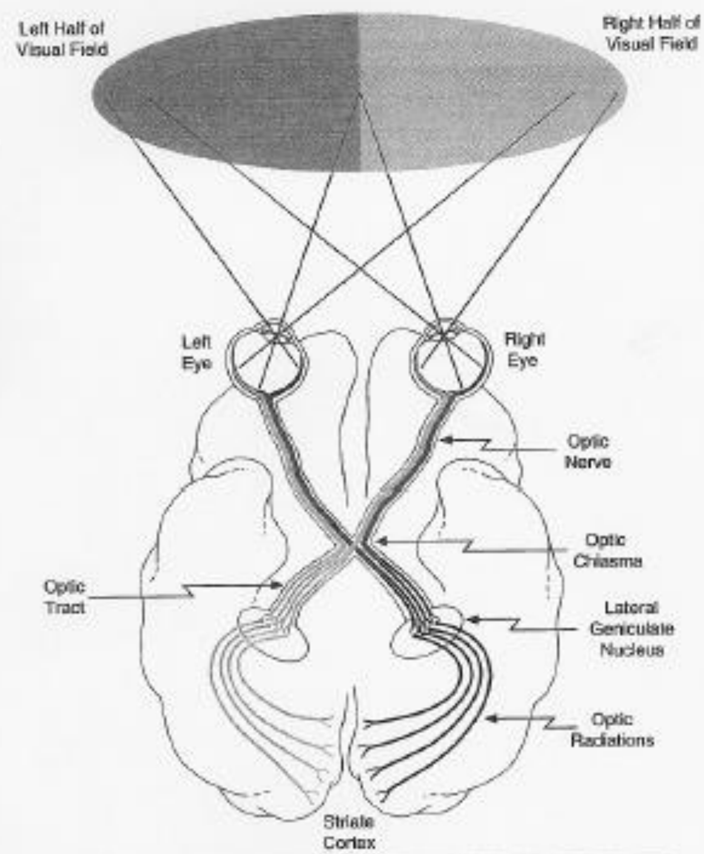


Figure 1.12 The major visual pathways from the eyes to the brain. The pattern of light striking each retina is encoded into nerve impulses, and these impulses are transmitted to the brain via the optic nerve that emerges from the retina. The left half of the visual field, which is imaged on the right half of each retina, is transmitted to the right half of the brain; the right half of the visual field, which is imaged on the left half of each retina, is transmitted to the left half of the brain. The cross-over of optic nerve fibers that is necessary to realize such a mapping takes place at the optic chiasma. From the optic chiasma, the nerve fibers proceed to the lateral geniculate nuclei via the optic tracts; from the lateral geniculate nuclei, nerve impulses are transmitted to the striate cortex of the brain via the optic radiations. It is important for binocular stereoscopic depth perception that each of the possibly two retinal images of a point in the visual field be mapped onto the same region of the brain.

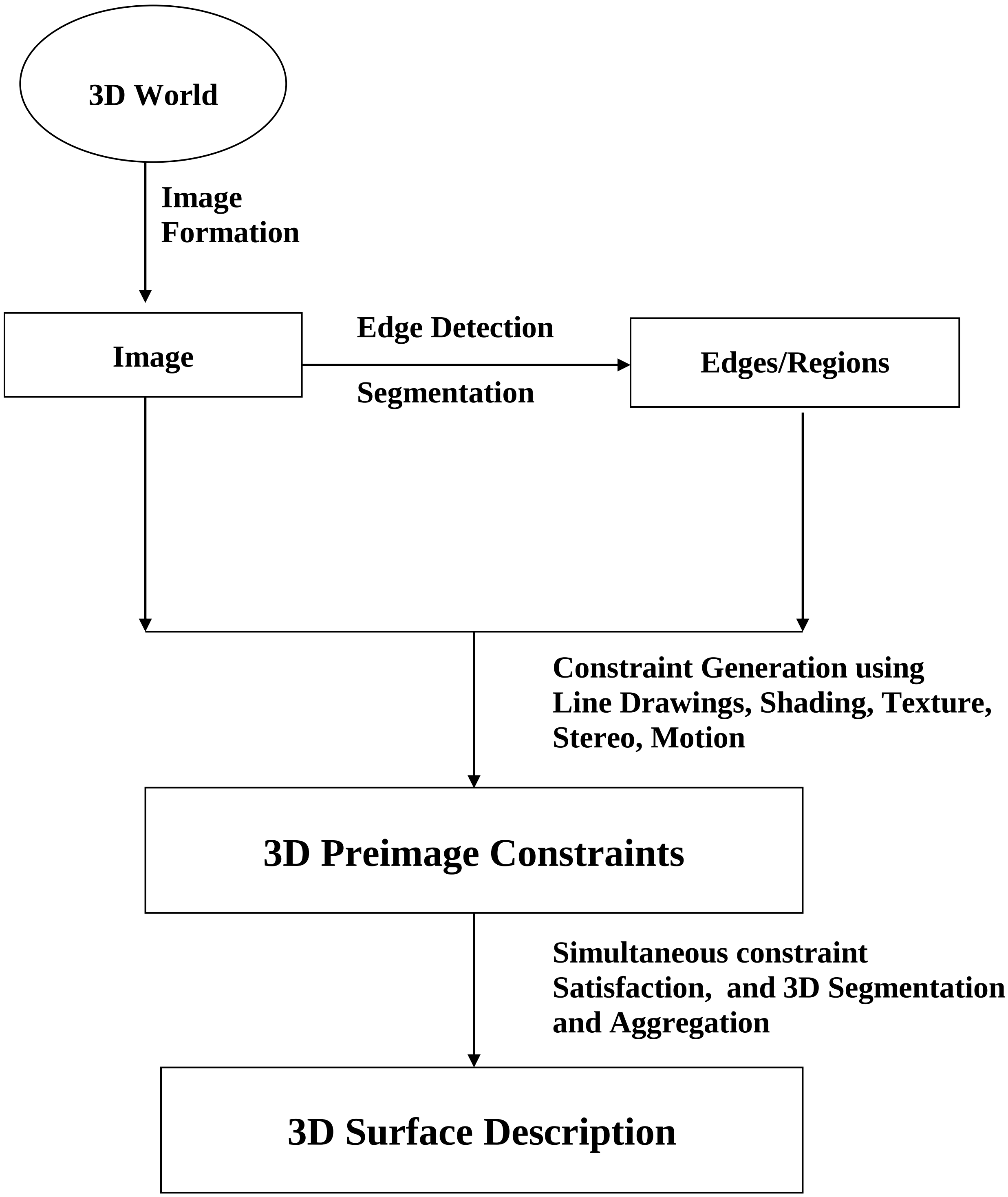
Left half of visual field which is imaged on the right half of each retina is transmitted to right half of brain. Vice versa for right half of visual field. From each eye emerges an optic nerve which carries electrical nervous signals from the eye to the brain. Fibers constituting each optic nerve can be divided into two groups: those that originate on the inner nasal side of the eye, and those that originate on its outer temporal side. Fibers originating on the temporal side of each eye go to the same side of the brain where they originate. In contrast fibers originating on each nasal side cross over at the ***optic chiasma*** and proceed to the opposite side of the brain as the eye where they originate. Left half of visual field is mapped onto right side of the brain. Right half of visual field is mapped onto left side of brain.

This is key:
The two retinal images of any point in the scene that is visible to both eyes are mapped onto the same region of the brain. It is the ***disparity*** between the two retinal images that makes *stereoscopic depth perception* possible within the field for view shared by the two eyes, also called the ***stereoscopic field***.

Try this experiment: Try to cap pen with one eye closed!

Optic tracts: Comprised of fibers that originate on same side of retina.
Lateral geniculate nucleus: Relay station to cortex.
Striate cortex: Visual center.

General Purpose Computer Vision



Use to Navigate, Recognize, Predict, Manipulate, etc.

2. IMAGE FORMATION

Image formation occurs when a sensor registers radiation that has interacted with physical objects. Mathematical model of imaging has several different components:

1. **Image function** = abstraction of image.
2. **Geometric model** = describes how three dimensions are projected into two.
3. **Radiometrical model** = shows how the imaging geometry, light sources, and reflectance properties of objects affect the light measurement at the sensor.
4. **Spatial frequency model** = describes how spatial variations of the image may be characterized in a transform domain.
5. **Color model** = describes how different spectral measurements are related to image colors.
6. **Digitizing model** = describes process of obtaining discrete samples.

2.1 Image Model

Image function is a mathematical representation of an image. Generally vector-valued, many times discrete. Images are typically presented by functions of two spatial variables

$f(\boldsymbol{x}) = f(x, y) =$ brightness of the gray level of the image at spatial coordinate (x, y).

Multispectral image

$\boldsymbol{f} = (f_1 ..., f_n)$

Key example is color image,

$\boldsymbol{f}(\boldsymbol{x}) = (f_{red}(\boldsymbol{x}), f_{blue}(\boldsymbol{x}), f_{green}(\boldsymbol{x}))$

Time-varying: $f(\boldsymbol{x}, t)$

Three-dimensional: $f(x, y, z)$

Important part of the formation process is the conversion of the image representation from a continuous function to a discrete function: ***delta function***. Basis of ***sampling***:

Definition:

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}$$

$$\int_{-\infty}^{\infty} \delta(x)\ dx = 1 \qquad \int_{-\infty}^{\infty} f(x)\ \delta(x-a)\ dx = f(a) \quad (\textbf{\textit{sifting property}})$$

2.2 Imaging Geometry

Pinhole camera:

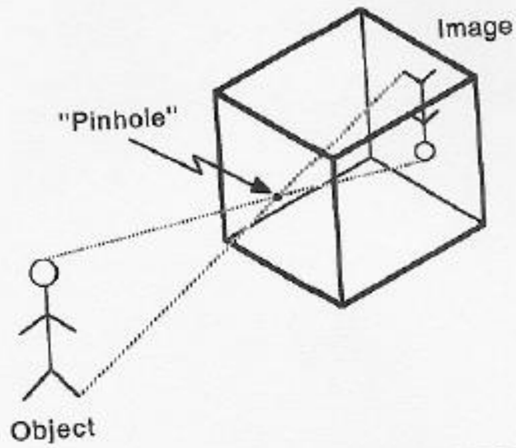


Figure 2.2 Pinhole camera. A pinhole camera is the simplest device to form an image of a three-dimensional scene on a two-dimensional surface: Straight rays of light pass through a "pinhole" and form an inverted image on a surface.

Image is reversed.

More intuitive to recompose the geometry so that the point of projection corresponds to a *viewpoint behind* the image plane, and the image is formed right side up.

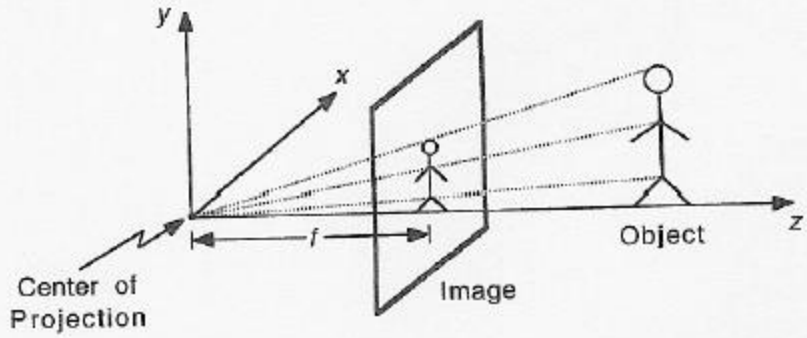


Figure 2.3 Perspective projection. In *perspective projection*, each object point is projected onto a surface along a straight line through a fixed point called the *center of projection*. (Throughout this book, we shall use the terms *object* and *scene* interchangeably.) The projection surface here is a plane. Perspective projection closely models the geometry of image formation in a pinhole camera (see Figure 2.2), except, in perspective projection, we are free to choose the location of the projection surface such that the image is not inverted.

I write down the coordinate system.

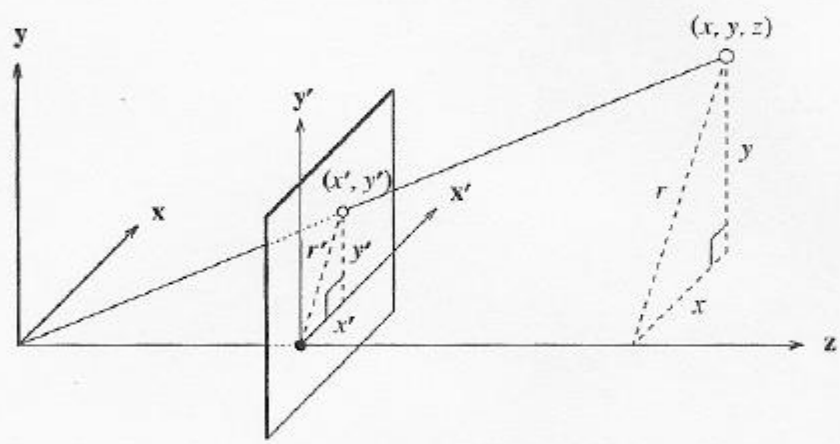


Figure 1.5: An illustration showing the line of sight that is used to calculate the projected point (x', y') from the object point (x, y, z).

Using similar triangles one can check:

$$\frac{y}{f-z}=\frac{y'}{f}\ ,\ \frac{x}{f-z}=\frac{x'}{f}$$

We have normalized image plane by taking $z=0$. We can generalize. We get general formula for image projection through pinhole camera:

Perspective transformation maps (x,y,z) to

$$\underbrace{(x',y',z')=\left(\frac{fx}{f-z}\ ,\ \frac{fy}{f-z}\ ,\ \frac{fz}{f-z}\right)}$$

Projective or orthographic projection

Note that as $z\to f$, coordinates $\to\infty$.

2.3 Reflectance

Physics of reflectance of objects: brightness, radiant intensity, irradiance. I won't discuss these concepts in detail. For those interested, see Horn, ***Robot Vision***.

2.4 Spatial Properties: Fourier Transform

An image is a spatially varying function. An important way to analyze spatial variations is the decomposition of an image function into a set of orthogonal functions, one such set being the Fourier (sinusoidal) functions. The ***Fourier transform*** may be used to transform the intensity image into the domain of *spatial frequency*. I will illustrate use *one-dimensional* Fourier transform. Easily generalized to any number of dimensions. I assume that you have already taken course in signal processing. The definition is:

$$\mathcal{F}\left[f(x)\right]\equiv F(u),$$

where

$$F(u)=\int_{-\infty}^{\infty} f(x)\,e^{-j2\pi ux}\,dx\,,\ j=\sqrt{-1}$$

Fourier transform has an inverse:

$$\mathcal{F}^{-1}\left[F(u)\right]=f(x),$$

with

$$f(x)=\int_{-\infty}^{\infty} F(u)\,e^{j2\pi ux}\,du\,.$$

The transform $F(u)$ is just another representation of image function. Choose point x_0, and note that

$$e^{j2\pi ux_0} = \cos 2\pi ux_0 + j \sin 2\pi ux_0$$

Then

$$f(x_0) = \int_{-\infty}^{\infty} F(u) \; [\cos 2\pi ux_0 + j \sin 2\pi ux_0] \; du$$

Then particular point in the image can be represented as a weighted sum of sinusoidal patterns ***weighted*** by $F(u)$. Low spatial frequencies account for “slowly” varying gray levels in an image, such as the variation of intensity over a continuous surface. High-frequency components are associated with “quickly varying” information, such as edges.

I have done continuous version. There is a ***discrete Fourier transform*** (DFT) and fast ways of computation ***fast Fourier transform*** (FFT).