Lecture 1
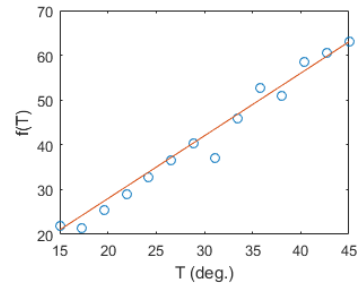### Fundamentals of Imaging
Lecture Notes by Assaf Tal

An Overview of In-Vivo Imaging

Imaging is Concerned With Mapping Meaningful (Physiological) Quantities Spatially

Imaging in general is concerned with mapping the spatial distribution of some parameter or set of parameters. For example, we might want to map the temperature T of a body as a function of position, T($\mathbf{r}$). This can be done using infrared cameras that image wavelength ~ 10 μm (these introduce questions of *penetration depth*, as IR radiation gets absorbed after a certain distance traveled in most matter).

One Rarely Maps The Quantity of Interest Directly

Rarely does one map a quantity directly, and most imaging is indirect: one maps a function *f* of the desired quantity, f(T($\mathbf{r}$)), and attempts to extract T out of it. An IR camera does not yield the temperature T directly. Rather, one measures radiation intensity in the infrared spectrum. The intensity depends on the body's temperature, and one needs to be able to connect the two: temperature $\leftrightarrow$ intensity. If one is lucky, there exists a known as simple relation between the two, but sometimes this is nontrivial. For example, the amount of IR energy, f(T($\mathbf{r}$)) might be proportional to the temperature: f(T)=αT, which would imply that, at each point $\mathbf{r}$ of our final image,

$$f\left(T\left(\mathbf{r}\right)\right) = \alpha T\left(\mathbf{r}\right)$$

To uncover α, one would probably carry out a set of calibration measurements by varying T and measuring f(T) (with noise!), and then carrying out a linear fit:

Of course, real life can be a lot more complicated, and there might not even exist a closed form expression (in MRI it usually does).

The Hallmarks of a Good Image: SNR, Resolution, Time

From an engineering point of view, the **signal to noise ratio** (SNR), **spatial resolution** and total **acquisition time** are the defining characteristics of an image. All three are usually adjusted to yield the best possible **contrast**.
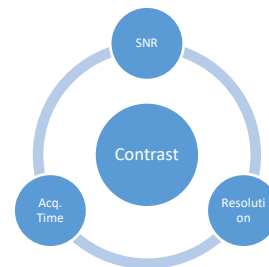
These are all very intuitive quantities. First, resolution:
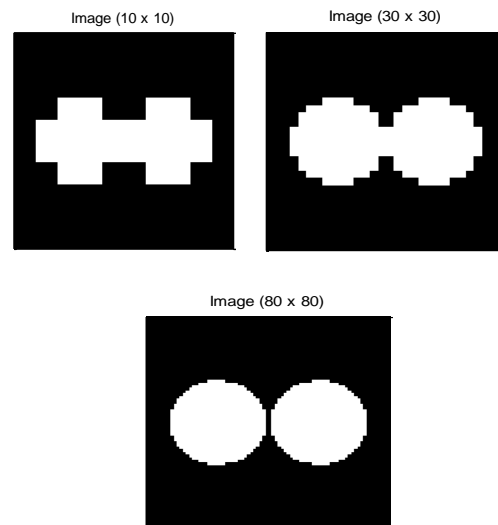
The images show the same object, two circles, with increasing image resolution (number of pixels). One talks of **pixels** when discussing 2D images, and **voxels** (=volumetric pixels) when discussing 3D images, and the **nominal spatial resolution**[1] is defined as the size of the pixel/voxel. The **matrix size** is the number of pixels or voxels, which for the images above is 10×10, 30×30 and 80×80.

The resolution could be different along different axes in space. An **isotropic voxel** is one having the same nominal resolution along all three spatial axes.

able to produce images at a much faster rate with a temporal resolution of ~10-100 ms.

The signal to noise is defined for a region, and is the ratio of the absolute value of the signal at a point divided by the standard deviation (SD) of the noise in the image. In many images, MRI included, noise is mostly random, so the noise's SD is calculated from an "empty" region containing no image and just noise. This is not always a perfect choice for reasons we'll encounter down the road, but it should give a rough estimate for the SNR.

Modality	Typical Spatial Resolution (Linear, in mm)	Typical Temporal Resolution	Picks up signals from ...
MRI: Anatomical (3D)	1	2-5 minutes	Water
MRI: Spectroscopy	10	~10 minutes	Small metabolites
MRI: BOLD-fMRI	2-3	~sec	Oxygenated blood
Positron Emission Tomography	1-5	20-60 minutes	Radiotracers (FDG, ^{82}Rb-Chloride, etc)
Ultrasound	1	~10-100 ms	Tissue acoustic impedance
Computerized Tomography (CT)	1	few minutes	Tissue radiation absorption
Screen film radiography ("X-ray")	0.1		Tissue radiation absorption
Magnetoencephalography (MEG)	1	0.1-1 ms	Magnetic fields generated by Electrical currents
Electroencephalography (EEG)	Problematic ...	0.1-1 ms	Electrical currents

Table 1. Some typical imaging modalities in humans and their resolutions

Pixel   Voxel   Isotropic voxel
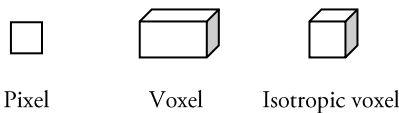
$$SNR = \frac{f(T)}{\sigma}$$

In general, one needs to specify the spatial resolution along each of the spatial axes of the voxel. Typical spatial resolutions of imaging modalities are shown in Table 1.

In many imaging modalities the **temporal resolution** is also important. For example, if we are interested in imaging the beating heart (~1 Hz, or ~ 1 beat per second), our imaging apparatus must be

A high SNR makes the image more discernable from the background. It also makes inverting $f(T)$ (to obtain T) easier. This is illustrated in the figure below.

[1] Most people omit the "nominal" when talking about number of pixels or voxels. Later on when we introduce

the point spread function we'll see that nominal resolution doesn't tell all the story.

SNR = 1.0   SNR = 1.0

SNR = 10.0   SNR = 10.0

SNR = 100.0 SNR = 100.0
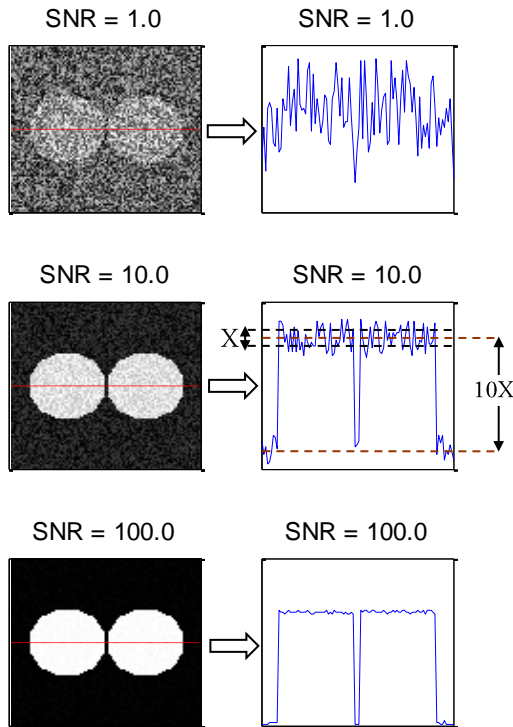
Finally, many systems require a considerable amount of time to acquire 2D or 3D images of a quantity of interest. A simple X-ray image takes less than a second to obtain, if one does not factor in the time needed to position the patient. An MRI scan can take seconds to minutes to tens of minutes. CT scans and ultrasound scans also take minutes to complete, and additional time to be deciphered by a specially trained radiologist. Radiology of the sub-specialty of medicine which is concerned with reading medical images.

Sometimes, Contrast Beats All

One often images a particular imaged quantity, T, not because it is interesting in itself but because it changes between, say, healthy and pathological tissue. Temperature, for example, is known to increase when there is infection. One is then interested in resolving the difference:

$$T_{\text{healthy}} - T_{\text{pathological}}$$

The ability to resolve this difference will depend on whether it is big or small in comparison to the "natural variation" in the image – that is, the noise. Therefore, the quantity we want to maximize is the

ratio of the two, known as the **Contrast to Noise Ratio** (CNR):

$$CNR = \frac{f\left(T_{\text{healthy}}\right) - f\left(T_{\text{pathological}}\right)}{\sigma}$$

A CNR>>1 is necessary for reliable differentiation between the two tissues.

The following figure shows two spheres with increasing CNR:

CNR = 1.0   CNR = 1.0

CNR = 5.0   CNR = 5.0

CNR = 10.0   CNR = 10.0
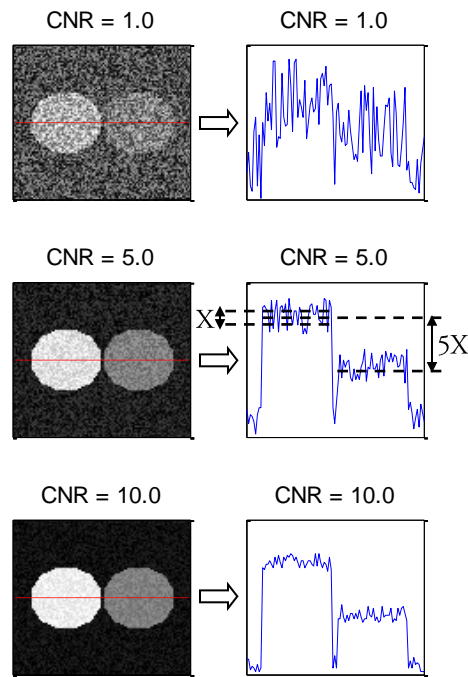
It is interesting to note that it is much easier for us to detect a difference between the spheres in the CNR=1.0 case when viewing the 2D image compared to the 1D slice through the image. This is due to multiple factors, some of which physical and some of which are rooted in neuropsychology and the way our brain and vision work (e.g. in the 2D case we're "integrating" over a larger region which gives us a better "effective CNR").

Quantities Are Averaged Over a Voxel

When one presents you with a map of some quantity T($\mathbf{r}$), you must keep in mind that the quantity is not really measured at an infinitely sharp point $\mathbf{r}$, but almost always is averaged somehow over the entire pixel (2D) or voxel (3D). Therefore, when we speak of "the temperature at the point $\mathbf{r}_0$,"

we really mean "the temperature averaged over a voxel centered at point r_0." One should always be careful when thinking of such quantities: if the voxel is much larger than the region in which temperature rises, the changes will be "washed out" by the averaging. This is known as a **partial volume effect**.

The Point Spread Function

Almost every imaging system can be analysed in terms of its **point spread function**, or PSF, which is the answer to the following question: if I put an infinitely narrow object in my imaging system (what physicists call a "delta function"), how will my final image look like? The resulting image is rarely an infinitely narrow object, but some finite, "blurred" version of it:


Input function: "delta" at **r**=0

Ouput: The point spread function
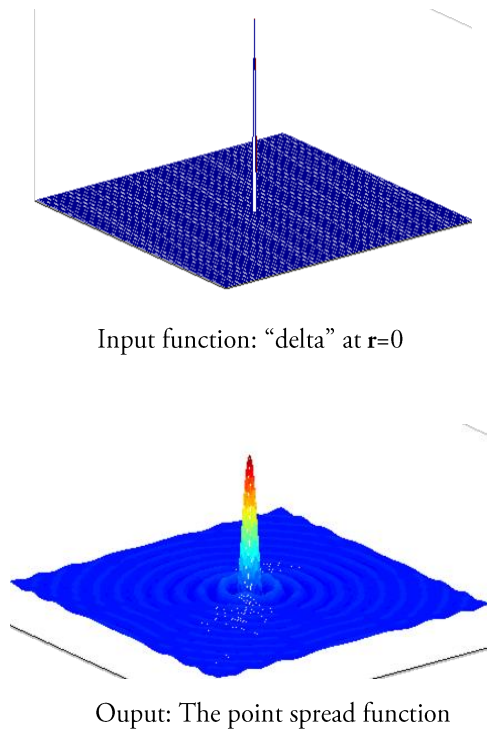
One can now treat a "real" object as made out of many infinitely narrow "delta functions." The final image is obtained by converting each "point" in the original image into the PSF, and adding everything up. This process is known as a convolution.

A **convolution** of an image I(**r**) with a point spread function PSF(**r**) is a transformation in which every point r_0 in I(**r**) is replaced by PSF(**r**-r_0). Mathematically, we denote this convolution as $(I \otimes PSF)(\mathbf{r})$ and define it as:

$$(I \otimes PSF)(\mathbf{r}) = \int_{\substack{all \\ space}} I(\mathbf{r}')\, PSF(\mathbf{r} - \mathbf{r}')\, d\mathbf{r}'$$

The way to imagine this is as follows: imagine the PSF "sliding along" the x-axis, with its center at x', while the image I(x') remains fixed.

Example: Suppose our 1D image is a rectangle of unit height and width L, and our PSF is a rectangle of unit height and width Z. The convolution remains zero as long as $x' < -\frac{Z}{2} - \frac{L}{2}$. As x' moves between $-\frac{Z}{2} - \frac{L}{2}$ and $+\frac{Z}{2} - \frac{L}{2}$, the amount of overlap between the image and the PSF increases linearly. The overlap – and hence the integral's value – becomes constant for $+\frac{Z}{2} - \frac{L}{2} < x' < \frac{L}{2} - \frac{Z}{2}$, and then drops linearly down to zero:
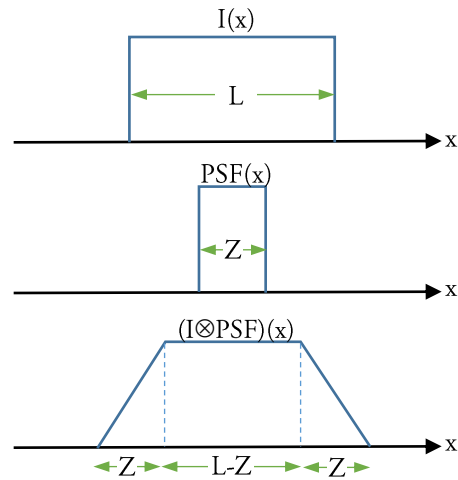

We've remarked that nominal resolution only tells half the story, and the PSF is the other half. Even if the nominal resolution is very high, you still might not be able to resolve structures if your PSF is too wide. For example, suppose that you have many detectors that are very finely spaced apart in 1 mm intervals. However, each detector isn't very good and it captures light from a wide range of

positions, about 1 cm wide. Even though you have many finely spaced detectors – your nominal resolution is high – you will still not be able to resolve structures that are closer than a centimeter, because they'll get blurred up by the PSF of the detectors:


Detectors   Object   Image
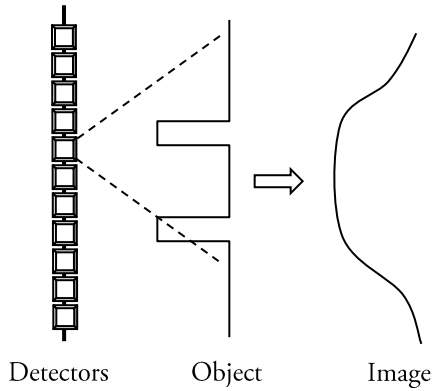
This is also illustrated in the following set of drawings which show what happens as we increase the spatial resolution of the two spheres presented in the previous sections, but convolved with the same unchanging PSF:
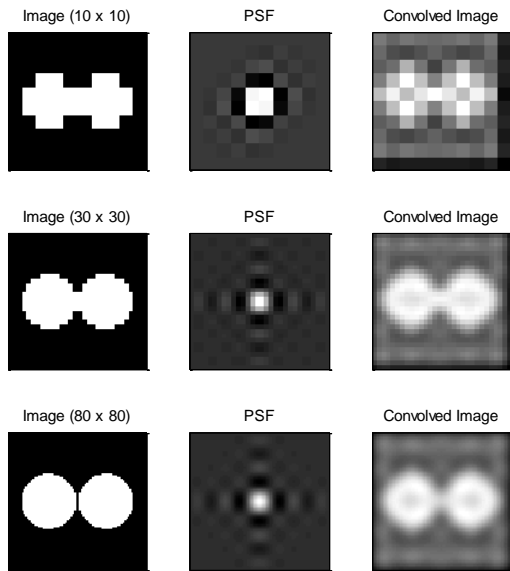

We see that past a certain point increasing the resolution does little to improve the final image. This is because we're sampling the convolved image, not the actual image!

The discussion focused on **spatial PSF**s which are easy to visualize. However, the imaging process also has a **temporal PSF**. A very famous example of this is obtained in blood oxygenation level dependent functional MRI (**BOLD-fMRI**) experiments, which measure hemodynamic changes in the brain in response to an external "task" or "stimulus". For example, a particular "input" is turned on - say, a particular flashing light – and an "MRI signal" is measured from a region of interest - say, the visual cortex. The visual cortex will almost always get activated in response to the flashing lights, meaning we will see a temporal change in the measured signal (the meaning of which we will discuss in-depth later on in the course). This signal usually looks as if it was convolved with a temporal PSF that has a width of several seconds, meaning that BOLD-fMRI has a point spread function a few seconds wide which looks like this:
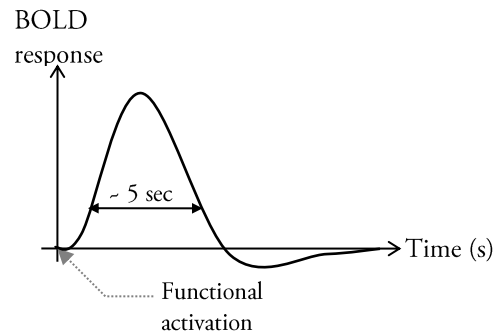
This means that, if we were to measure the response of a brain region to an external stimulus, we could predict it by convoluting the stimulus signal with the PSF:
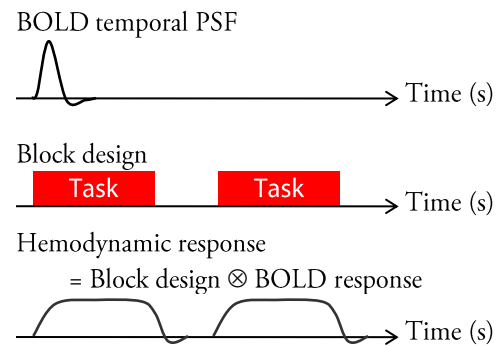
The brain however is not truly a linear response and thinking in terms of point spread functions and simple convolutions fails when tasks are spaced too closely together.

Image Orientation Names

In general anatomy, several terms describing "direction" appear repeatedly:

- Rostral (towards the nose).
- Caudal (towards the tail).
- Cranial (towards the skull).
- Dorsal (towards the back).
- Ventral (towards the belly, or underside).
- Posterior (towards the rear/tail).
- Anterior (towards the head/front).
- Left and right, which are self-explanatory.
- Medial (towards the center of the body).
- Lateral (towards the edge of the body).
- Inferior (towards the bottom of the body).
- Superior (towards the top of the body).

These terms are used to define several plane orientations:

- **Coronal** planes are perpendicular to the posterior-anterior axis[2].
- **Transverse** (or axial) planes are perpendicular to the caudal-cranial line.
- **Sagittal** planes are perpendicular to the left-right line.

For humans, they have the following interpretation:

Sagittal

Axial   Coronal
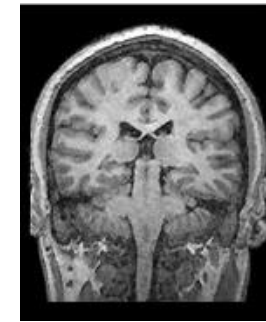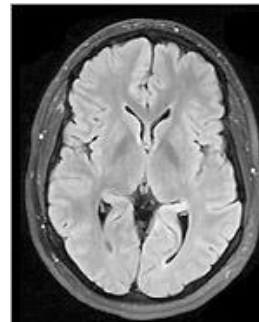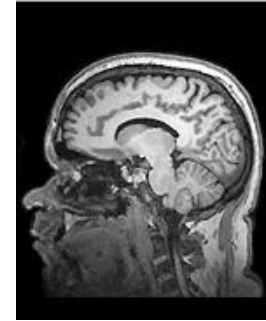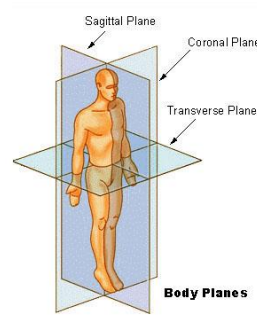
However, it is important to keep in mind things change a bit in animals. For example, the dorsal-ventral and the anterior-posterior directions are the same for humans, but different for four-legged animals or fish:
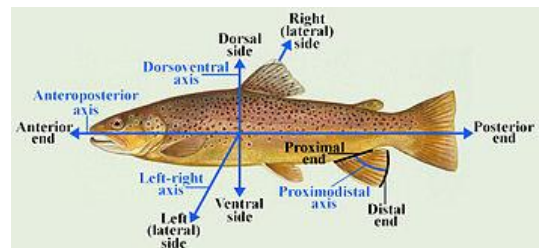

This is why the fin on the back of a shark is often called the dorsal fin. Thus, a coronal plane in a fish is "the same" as a transverse plane in humans. Two other examples:

1. The superior direction in humans is the head. The superior direction in animals is the back.

---

[2] There is some confusion related to coronal planes, since some sources define them as being perpendicular to the dorsal-ventral line – that is, dividing the body into "belly" and "back" sections. In humans, these are one and the same. However, in mice they are not! Imagine a mouse standing on four legs: The axis from

the belly to the back is actually perpendicular itself to the axis from the front to bottom. There is no clear cut "winner", although it does seem that the "front to back" definition seems to win out more if you look at scientific imaging literature in mice.

2. The rostral-caudal line in humans is curved, while in animals it is straight:
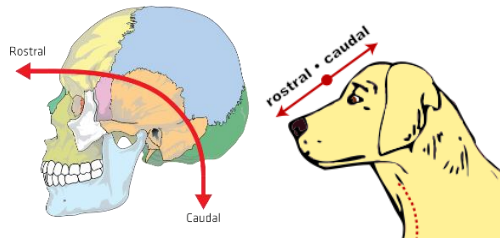
Another example of the sort of confusion these definitions create comes from mice.

The plot thickens somewhat more in the brain, where the **meaning** and not the **direction** of certain terms should be used. As an example, the prefrontal cortex is the part of the cortex in the brain which is close to the person's face. It is divided into "dorsal" and "ventral" parts:
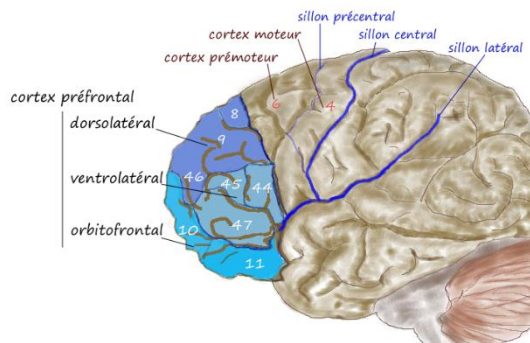
Here, the dorsal part is "above" the ventral part. This makes sense if you recall the meaning of the words: ventral means towards the "belly", or "underside" of the brain. Thus, for a human, one can speak of the "dorso-ventral line" connecting the actual back and belly, and can equally speak of "dorsal" and "ventral" divisions of the brain. The two uses involve different spatial directions – that is, unless the person is getting a massage and has his/her chin resting on the massage bed!