Diffusion

1 Introduction

We're now going to talk about diffusion processes. This is a physical process in which matter, or heat, for example, spread throughout a medium. This might seem like a non-sequitor with no connection to vision, but diffusion processes are relevant to segmentation for two reasons. First of all, diffusion is a kind of smoothing, and this gives us another set of tools to understand smoothing. For example, if you have a cup of hot coffee and you pour some cold milk into it, what happens to the heat? At first, parts of the cup are hot, and parts are cold. Over time, the heat spreads out, or is smoothed. Second, the result of diffusion can be found by solving a partial differential equation. So discussing diffusion will lead to a discussion of numerical methods for solving PDEs, which are going to be important throughout this course.

Here's what we want to accomplish today:

- Describe diffusion as a PDE.
- Describe an explicit method for solving this numerically.
- Point out that this explicit method is equivalent to repeated convolution with a small filter. It turns out that this is equivalent to convolution with a Gaussian. So diffusion low-pass filters a signal. Diffusion and Gaussian convolution and low-pass filtering are all equivalent.
- Point out that there are other ways to solve these equations, which have some useful properties (implicit method).

2 Diffusion as a PDE

One way to think about diffusion of a material in a medium is to imagine the world is divided up into tiny buckets. Each bucket has a discrete number of molecules, say. At every time step, a molecule can stay where it is or jump to a neighboring bucket. We make this simple by assuming it is:

homogenous: This means that the same thing happens everywhere. A particle has the same probability of doing things no matter which bucket it's in.

isotropic: This means that the particle is equally likely to jump to the left and right.

Some intuitions. If one bucket has a lot of particles, and its neighbor has few, then more particles jump from the first to the second than from the second to the first. So the two buckets tend to become more alike over time (smoothing).

We can take the limit of this model, as our buckets get smaller and our time steps get shorter. It is then natural to express the way in which the concentration of particles change with a differential equation.

$C(x,t)$ is the concentration of particles at position x at time t. When we are considering a fixed point in time, we may just write: C_x.

$J(x,t)$ represents the flux of particles. This is the rate at which particles move in the positive direction across the position x at time t.

$$J(x,t) = -D\frac{\partial C}{\partial x}. \tag{1}$$

This means that the flux is proportional to the difference in the concentration of material. If $\frac{\partial C}{\partial x}$ is big, it means the concentration is increasing, so more particles are flowing backwards than forwards. D is the diffusivity. This is a constant which indicates how rapidly particles diffuse in this medium. Note that D is a constant because the diffusion process is homogenous and isotropic. Otherwise, it could be a function of x.

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial x} \tag{2}$$

Change in concentration is inversely proportional to the change in flux. For example, if flux is constant, no change in concentration. This equation is basically the limit of having a small cell. The difference in flux on the two sides of the cell is the difference in how many particles are entering from the left, and how many are leaving to the right. Sign is negative because flux on the left side of the cell indicates how many particles are entering. If flux is dropping, more particles enter on the left than leave on the right.

Taking the partials of (1) w.r.t x we get: $\frac{\partial J}{\partial x} = -D\frac{\partial^2 C}{\partial x^2}$. Substituting into (2) we get:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{3}$$

This tells us that concentration changes in the direction of the second derivative. This tells us that the concentration gets smoother. For example, at a local maxima, the concentration drops, and at a local minima, the concentration increases.

This diffusion equation is also used in vision to smooth curves.

3 Numerical solution: Explicit method

We want to compute how C evolves over time. To do this numerically, the main issue is to take derivatives numerically. These will, of course, be an approximation. Taking a Taylor series expansion we find:

$$C_{i\pm 1} = C_i \pm \delta x\frac{\partial C}{\partial x} + \frac{1}{2}\delta x^2\frac{\partial^2 C}{\partial x^2} + O(\delta x^3).$$

Using this equation, and ignoring higher order terms we find that we can approximate

$$\frac{\partial C}{\partial x} = \frac{C_{i+1} - C_i}{\delta x} = \frac{C_i - C_{i-1}}{\delta x} = \frac{C_{i+1} - C_{i-1}}{2\delta x}$$

Through similar reasoning, we find:

$$\frac{\partial^2 C}{\partial x^2} \approx \frac{(C_{i+1} - C_i) - (C_i - C_{i-1})}{\delta x^2}$$

This is intuitive, since it's the difference in the derivative, approximated at C_i and C_{i+1}, divided by δx.

From (3) we get:

$$\frac{\partial C}{\partial t} \approx \frac{D}{\delta x^2}((C_{i+1} - C_i) - (C_i - C_{i-1}))$$

If we take another Taylor series expansion, we can compute the partial with respect to t in the same way, by taking the difference between C evaluated at two different times, and dividing by δt. So we have:

$$\frac{\partial C}{\partial t} \approx \frac{D}{\delta x^2}((C_{i+1} - C_i) - (C_i - C_{i-1})) \approx \frac{C(i, n+1) - C(i, n)}{\delta t}$$

$$C(i, n+1) \approx C(i, n) + \frac{D\delta t}{\delta x^2}(C(i+1, n) + C(i-1, n) - 2C(i, n))$$

Notice that if we set $\lambda = \frac{D\delta t}{deltax^2}$ this has the form:

$$C(i, n+1) = (1 - 2\lambda)C(i, n) + \lambda C(i+1, n) + \lambda C(i-1, n)$$

This is convolution with a filter that looks like

λ	$1 - 2\lambda$	λ

So, we can compute one time step of the diffusion equation with a convolution. To compute many time steps, we convolve with this many times. Due to the associative law, this is equivalent to convolving with one big Gaussian.

To see why this is, lets think about what happens to a single particle that is diffusing. At each tiny time step, it moves left or right. Let's define a random variable, x_i. x_i is $-\delta x$ if the particle moves left at time i, it's δx if it moves right, and 0 otherwise. So after T time steps the position of the particle is:

$$P = \sum x_i.$$

By the law of large numbers, P is a random variable with 0 mean and a Gaussian distribution. So after T time steps, the particle's position is a Gaussian. We can compute this by convolving its initial position with a Gaussian. Or, we've seen, we can do it with repeated convolution by a small filter, which is the same thing.

This means that convolving with a Gaussian tells us the solution to the diffusion equation after a fixed amount of time. This is the same as low pass filtering an image. So smoothing, low pass filtering, diffusion, all mean the same thing.