

Wavelets

We've been talking about multiscale as a way to build up descriptions of images that can correspond to segmentations. We're going to be talking about scale in the context of texture, because this is a way to try to get away from simple descriptions of images for segmentation that are based on pairwise relations between pixels. Texture is something that happens on a larger scale (it doesn't really make sense to talk about the texture of a pixel, just the texture of a region). So we need ways of describing images at some larger scale. Wavelets will be important to building descriptions of textures because they capture the frequency information of textures at local scales. Also, they are generally very important, and fit well with other multiscale ideas we've been discussing.

We've discussed two ways of representing images. Pixels are spatially localized, but they don't capture frequency information. The Fourier transform captures frequency information, but it is not spatially localized. We want something that is localized in space and frequency.

Windowed Fourier Transform

Multiply the signal by a windowing function, and then take the Fourier transform. This window slides over the signal, so a 1D signal is described with 2 degrees of freedom, translation and frequency. One way to do this is by multiplying a harmonic by a Gaussian, and then convolving with that. We note that this is spatially localized because of the Gaussian. Let's look at this by taking a Fourier transform.

$$w(x) = h(x)g(x)$$

w is a windowed harmonic, where h is a harmonic function (sin or cos of some frequency) and g is a Gaussian. Denoting Fourier transforms with capitals, we have, from the convolution theorem:

$$W(u) = H(u)*G(u),$$

where u is frequency. The transform of a Gaussian is a Gaussian, the transform of a harmonic is a delta function. So convolving a Gaussian with a delta just shifts it. This means the windowed harmonic is localized in space and frequency. And just like a Gaussian, the more localized it is in one domain, the less localized it is in the other domain.

This is called a Gabor function, and is in some sense the function that is optimally localized in space and time. The Heisenberg uncertainty principle says that it is impossible to have a function that is perfectly localized in space and time. Gabor functions are used as a model of neurons in the early visual system.

We can represent an image by convolving it with Gabors of different location and frequency. One disadvantage of this simple scheme is that we are fixing the scale of the

Gaussian, but we want to capture image events that occur at different scales. The intuition is that high frequencies can occur at very fine scales, but low frequency events occur at coarser scales. So we vary the scale together with the frequency.

Wavelets

Wavelets will scale as the frequency changes. We do this with a function, Ψ , the mother wavelet. Ψ integrates to 0 and has a norm of 1. Then we can translate $\Psi(t)$ by u by taking $\Psi(t-u)$. We scale by taking $\Psi^{u,s} = 1/\sqrt{s} \Psi((t-u)/s)$. The $1/\sqrt{s}$ is there so that the function still integrates to 1.

As an example, we will use the Haar wavelet, which is not necessarily the best, wavelet, but is certainly the simplest.

$$\begin{aligned} \Psi(t) &= -1 \text{ if } 0 \leq t \leq 1/2 \\ \Psi(t) &= 1 \text{ if } 1/2 \leq t \leq 1 \end{aligned}$$

Then, for example, $1/\sqrt{s} \Psi(t-u/s)$, with $s = 2$, is:

$$\begin{aligned} \Psi(t) &= -1/\sqrt{2} \text{ } 0 \leq t \leq 1 \\ \Psi(t) &= 1/\sqrt{2} \text{ } 1 \leq t \leq 2, \end{aligned}$$

We can see that this still has a norm of 1 because Ψ squared is $1/2$ between 0 and 2.

Wavelet Transform

This is the continuous case, sort of like the fourier transform. For all u, s we take:

$$\langle f, \Psi^{u,s} \rangle$$

This represents a 1D function f with a 2D set of values. This redundancy occurs because the wavelets are not orthogonal, so they form a redundant basis. One idea is that this will give us a big response for scales and locations where there is a lot of energy, ie., a localized frequency. There is also a nice theorem that states that:

$$f(t) = k \iint \langle f, \Psi^{u,s} \rangle \Psi^{u,s} du \frac{ds}{s^2}$$

where k is some constant. This is very much like the formula used to reconstruct a signal from its Fourier transform, and is like saying that a function is a linear combination of harmonics, in the fourier transform.

However, there is one big difference. In the fourier transform, the harmonics form an orthonormal basis. This is not the case with the wavelet transform. This is because even though two cosine functions may be quite similar, since they are infinite, they ultimately

go through all states of being in and out of phase. Since wavelets are spatially localized, this is not the case. This is particularly easy to see with the Haar wavelet.

Orthonormal Basis of Wavelets

So, it is useful to construct an orthonormal basis of wavelets. To do this, we don't take the wavelet at all scales and translations, but just at a discrete set of them that are orthonormal. The idea is that we shift each wavelet enough so that it is orthogonal to the others, and scale it enough so that it is orthogonal.

Example: Haar wavelet. We translate it in units of 1, and scale it by factors of 2. It is easy to verify that these are orthonormal.

Next, we need to show that this can be a basis for all functions. We will do this in a way that also demonstrates how we can efficiently take the transform. Our strategy is to isolate the high frequency information, and represent it with an orthonormal basis of fine scale wavelets. Then we can repeat this at coarser scales. But we can also shrink the signal, so that performing this at coarse scales is more efficient. This should sound familiar.

Suppose S is an arbitrary, reasonable function. We want to show that we can approximate it arbitrarily well as a linear combination of Haar wavelets. First, we can approximate f arbitrarily well with a piecewise constant function where the pieces are all the same width, k . This approximation is the one we use to take the Riemann integral, and we have similar requirements on the function f .

Next, we're going to begin our approximation with Haar wavelets that have a width of $2k$. Let's consider the first part of f , assuming these two pieces have heights of a and b . If we take the inner product of this with a Haar in the same position we have $(b-a)/2$. So that means we're approximating that part of the signal with a function that is:

$-(b-a)/2$ and then $(b-a)/2$.

If we subtract this from f , we get

$(b+a)/2$ and then $(b+a)/2$.

So the residual function is constant over a range with a width of $2k$.

This means that by projecting f onto the Haar's at the finest level, we're left with a function that is one level coarser. At some point, we're left with a constant function. If we keep going, we get a bigger and bigger constant function that has a smaller and smaller amplitude. The sum of the square of this converges to 0.

Notice that when we move to a coarser scale, we could just shrink the whole function by a factor of 2. This means the work we do at each scale is half that at the previous scale. This should sound familiar.

Some other points

It is important to know that there are many types of wavelets, which have more desirable properties than Haars. First, the fact that the Haar wavelet isn't smooth is often a bad thing. We can get smoother wavelets. Second, the Haar is maximally localized in space. This makes intuitive sense, since it drops to zero so quickly. But this means that it is minimally localized in frequency, compared to other wavelets. Often, we would want more of a compromise.

Even if a set of wavelets is not orthonormal, it can still be a *frame*, which has many desirable properties. If wavelets form a frame, we can reconstruct a signal from its projection onto the frame. This is analogous to saying that the wavelets form a redundant basis.