# Scale spaces for $I(x, y)$

Today we generalize the scale space ideas from 1D images $I(x)$ to 2D images $I(x, y)$. You should read these notes together with the slides. (This is true for most lectures. It is especially true today.)

## Gradients and the second moment matrix

If we have a 2D image, then we can define a gradient filter,

$$\nabla g_\sigma(x, y) = (\frac{\partial g_\sigma}{\partial x}(x, y), \frac{\partial g_\sigma}{\partial y}(x, y)).$$

Similar to the 1D case, the response of filter $\nabla g_\sigma(x, y)$ to a vertical edge

$$I(x, y) = au(x - x_0)$$

will be independent of $\sigma$ along the edge, i.e. the line $x = x_0$. (By contrast, the response of a gradient filter $\nabla G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}\nabla g_\sigma(x, y)$ to an edge would depend on $\sigma$ at the location of the edge. The argument is exactly the same as in the 1D case.) Similarly, the response of filter $\nabla g_\sigma(x, y)$ to a horizontal edge $I(x, y) = au(y - y_0)$ will be independent of $\sigma$ along the edge.

For a general image $I(x, y)$, we can define a scale space from these gradients, namely

$$\nabla g_\sigma(x, y) * I(x, y) = (\frac{\partial g_\sigma}{\partial x}, \frac{\partial g_\sigma}{\partial y}) * I(x, y).$$

If $I(x, y)$ is an edge at arbitrary orientation, then again the magnitude of the gradient will be independent of $\sigma$ along the edge. The magnitude of the gradient will be a Gaussian shaped ridge in the neighborhood of the edge.

Next, recall the second moment matrix

$$\mathbf{M} = \sum_{(x,y)\in Ngd(x_0,y_0)} \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & (\frac{\partial I}{\partial y})^2 \end{bmatrix}$$

where the symbols $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ hide the fact that we have blurred, prior to taking the derivative. Specifically, we suppose that each of the partial derivatives is computed by $\nabla g_\sigma(x, y) * I(x, y)$. Using a range of $\sigma$ values gives us a scale space of 2nd moment matrices, namely a family of second moment matrices that are defined over the domain $(x, y, \sigma)$.

Observe that there is another scale here, namely the summation neighborhood $Ngd$. If we are using a larger $\sigma$ *e.g.* to blur away the noise, then we want to use a larger neighborhood size as well. The common strategy for doing this is to scale the width and height of $Ngd$ by some multiple of the $\sigma$ used to blur away the noise.

One other point worth mentioning is that it is common to define the the second moment matrix in a slightly different way, such that pixels that are farther from the neighborhood center, $(x_0, y_0)$, will contribute less. This weighting is justified by the fact that the linear approximation (Taylor series) that was used to construct the second moment matrix holds best for points near $(x_0, y_0)$.

One commonly uses a weighting function, $g_{\sigma_I}(x, y, \sigma_I)$, where $\sigma_I$ is called the "integration scale". One can then write

$$\begin{aligned}\mathbf{M}(x,y) &= \sum_{(x',y')} g_{\sigma_I}(x-x', y-y') \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & (\frac{\partial I}{\partial y})^2 \end{bmatrix} \\ &= g_{\sigma_I}(x,y) * \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & (\frac{\partial I}{\partial y})^2 \end{bmatrix}.\end{aligned}$$

Note that there are two "$g$" functions here. There is the $g_{\sigma_I}(x, y)$ used to integrate the components of the second moment matrix over a neighborhood, and there is the $g_{\sigma_D}(x, y)$ which is used to blur the noisy image prior to taking the gradient. The latter is not written explicitly above, but it there. The $I$ and $D$ refer to the *integration scale* and *derivative* scale, respectively, and are sometimes called the *outer scale* and *inner scale*. Typically the ratio $\sigma_I : \sigma_D$ is constant, for example, 3.

In class, I gave several examples of gradient images and Harris corners that were detected. See the slides!

## Bars

Last lecture we discussed box detection for 1d images. Recall that if we filter an edge $u(x-x_0)$ with $\sigma \frac{\partial^2}{\partial x^2} g_\sigma(x)$, then the peak response will be independent of $\sigma$ and will occur at $x_0 \pm \sigma$. Equivalently, since $g_\sigma(x) = \sqrt{2\pi}\sigma G(x, \sigma)$, if we filter an edge $u(x - x_0)$ with $\sigma^2 G(x, \sigma)$, then the peak response will be independent of $\sigma$ and will occur at $x_0 \pm \sigma$. We would like to extend this sort of property to 2D.

First, consider a 2D vertical edge $I(x, y) = u(x - x_0)$. If we blur this edge with a 2D Gaussian $G(x, y, \sigma)$ then (you can verify)

$$I(x, y) * G(x, y, \sigma) = u(x - x_0) * G(x, \sigma)$$

where we have a 2D convolution on the left and a 1D convolution on the right, which is allowed because the function on the left does *not* depend on $y$. In particular, note that on the left we get the same blur profile (in the $x$ direction) as we do on the right.

If we take the second derivative in $x$, and multiply by $\sigma^2$, we get

$$\begin{aligned}\sigma^2 \frac{\partial^2}{\partial x^2} u(x-x_0) * G(x,y,\sigma) &= \;= \sigma^2 \frac{\partial^2}{\partial x^2} u(x-x_0) * G(x,\sigma) \\ &= \frac{1}{\sqrt{2\pi}} \sigma \frac{\partial^2}{\partial x^2} u(x-x_0) * g_\sigma(x)\end{aligned}$$

It follows that the peak response in the 2D case occurs along the two vertical lines $x = x_0 \pm \sigma$, and this peak response is independent of $\sigma$. Obviously, a similar argument would hold for a horizontal edge $u(y - y_0)$, namely filtering with $\sigma^2 \frac{\partial^2}{\partial y^2} G(x, y, \sigma)$ would produce a peak response at $y = y_0 \pm \sigma$ and this peak response would be independent of $\sigma$.

Next, consider a vertical *bar* image,

$$I(x, y) = u(x + \sigma_0) - u(x - \sigma_0).$$

Following the same argument as above, along with the observations from last class, if we filter the bar image with $G(x, y, \sigma)$ and then take the second derivative in the $x$ direction, and multiply by $\sigma^2$, we get a peak response at $x = x_0$ which is the line at the center of the bar, and this peak response would occur at $(x, y, \sigma) = (x_0, y, \sigma_0)$. In particular, the peak response would occur at the scale that corresponds to the halfwidth of the bar.

Exactly the same argument holds for bars in the $y$ direction, of course. To detect a bar in the $y$ direction we would filter with $\sigma^2 \frac{\partial^2}{\partial y^2} G(x, y, \sigma)$.

## Laplacian of a Gaussian

How do we detect bars in arbitrary directions? Consider the "Laplacian operator" which is defined to be the sum $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. The Laplacian of a Gaussian is **[ Eq. (*) below: modified Nov. 23, 2010]**

$$\begin{aligned} \nabla^2 G(x, y, \sigma) &= \frac{\partial^2 G}{\partial x^2}(x, y, \sigma) + \frac{\partial^2 G}{\partial y^2}(x, y, \sigma) \\ &= \frac{1}{\pi \sigma^4}\left(\frac{x^2 + y^2}{2\sigma^2} - 1\right) e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (*) \end{aligned}$$

Note that the $\nabla^2 G(x, y, \sigma)$ is radially symmetric, i.e. it depends only on $x^2 + y^2$. Thus, it will give the same response along a vertical bar vs. a horizontal bar vs. an arbitrarily oriented bar.

Given an image $I(x, y)$, one can define a Laplacian of a Gaussian scale space, $I(x, y) * \nabla^2 G(x, y, \sigma)$. For a vertical bar, the second derivative in the $y$ direction is zero, and so the response is the same as we saw in the 1D case from last lecture. (Similarly, for a horizontal bar, the second derivative in the $x$ direction is zero.) In particular, the filter $\sigma^2 \nabla^2 G(x, y, \sigma)$ will produce a peak response along the midline of the bar image $I(x, y)$, regardless of the orientation of the bar, and the scale of the peak will occur at the $\sigma = \sigma_0$. Because the Laplacian is radially symmetric, this scale space property holds for a bar of any orientation.

Finally, note that in the 2D case,

$$g_\sigma(x, y) = 2\pi\sigma^2 G(x, y, \sigma)$$

since the 2D Gaussian is a product of two 1D Gaussians. Thus, we could alternatively use the 2D filter $\nabla^2 g_\sigma(x, y)$ to detect the position and scale of a bar.

## 2D Box detection

We can define a 2D box function by taking the product of two 1D box functions,

$$I(x, y) = (u(x + \sigma_0) - u(x - \sigma_0))(u(y + \sigma_0) - u(y - \sigma_0)).$$

This function has a square shape. Following the same arguments as for bars, it is easy to see that the scale space

$$I(x, y, \sigma) = I(x, y) * \sigma^2 \nabla^2 G(x, y, \sigma)$$

has a peak value at $(x, y, \sigma) = (x_0, y_0, \sigma_0)$. Thus, to find boxes, one can look for local maxima or minima in this scale space. See the slides for examples.