Image processing

Brian Curless
CSE 557
Autumn 2015

Reading

Jain, Kasturi, Schunck, *Machine Vision*. McGraw-Hill, 1995. Sections 4.2-4.4, 4.5(intro), 4.5.5, 4.5.6, 5.1-5.4. [online handout]

What is an image?

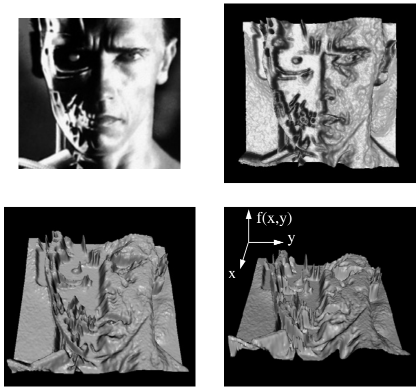
We can think of an **image** as a function, *f*, from $\mathbb{R}^2$ to $\mathbb{R}$:

- $f(x, y)$ gives the intensity of a channel at position (x, y)
- Realistically, we expect the image only to be defined over a rectangle, with a finite range:
  - $f: [a, b] \times [c, d] \rightarrow [0,1]$

A color image is just three functions pasted together. We can write this as a "vector-valued" function:

$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

Images as functions



What is a digital image?

In computer graphics, we usually operate on **digital** (**discrete**) images:

- **Sample** the space on a regular grid
- **Quantize** each sample (round to nearest integer)

If our samples are Δ apart, we can write this as:

$$f[i,j] = \text{Quantize}\{\, f(i\Delta, j\Delta) \,\}$$

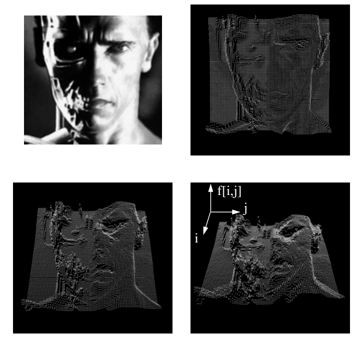


Image processing

An **image processing** operation typically defines a new image g in terms of an existing image f.

The simplest operations are those that transform each pixel in isolation. These pixel-to-pixel operations can be written:

$$g(x,y) = t(f(x,y))$$

Examples: threshold, RGB → grayscale

Note: a typical choice for mapping to grayscale is to apply the YIQ television matrix and keep the Y.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

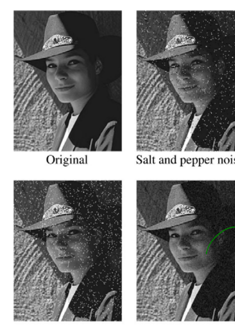
Noise

Image processing is also useful for noise reduction and edge enhancement. We will focus on these applications for the remainder of the lecture…

Poisson shot noise

$\sigma^2 \sim I = \mu$

$SNR = \frac{\mu}{\sigma} = \frac{I}{\sqrt{I}} = \sqrt{I}$



Original | Salt and pepper noise

Impulse noise | Gaussian noise

$I \sim \mathcal{N}(\mu, \sigma)$

$\approx I_{true} + \mathcal{N}(0, \sigma)$

Common types of noise:

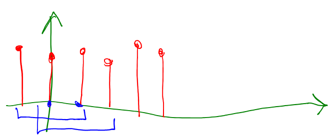
- **Salt and pepper noise**: contains random occurrences of black and white pixels
- **Impulse noise:** contains random occurrences of white pixels
- **Gaussian noise**: variations in intensity drawn from a Gaussian normal distribution

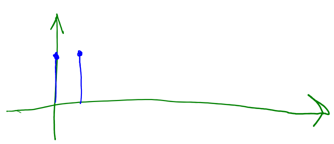
Ideal noise reduction

Ideal noise reduction

Practical noise reduction

How can we “smooth” away noise in a single image?





Is there a more abstract way to represent this sort of operation? *Of course there is!*

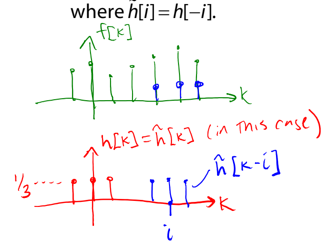
Discrete convolution

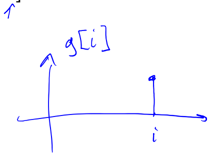
One of the most common methods for filtering an image is called **discrete convolution**. (We will just call this “convolution” from here on.)

In 1D, convolution is defined as:

$$\begin{aligned} g[i] &= f[i] * h[i] \\ &= \sum_k f[k]h[i-k] \\ &= \sum_k f[k]\tilde{h}[k-i] \end{aligned}$$

where $\tilde{h}[i] = h[-i]$.





“Flipping” the kernel (i.e., working with $h[-i]$) is mathematically important. In practice, though, you can assume kernels are pre-flipped unless I say otherwise.

Convolution in 2D

In two dimensions, convolution becomes:

$$\begin{aligned} g[i,j] &= f[i,j] * h[i,j] \\ &= \sum_{\ell}\sum_{k} f[k,\ell]\,h[i-k,j-\ell] \\ &= \sum_{\ell}\sum_{k} f[k,\ell]\,\tilde{h}[k-i,\ell-j] \end{aligned}$$

where $\tilde{h}[i,j] = h[-i,-j]$.

Again, "flipping" the kernel (i.e., working with $h[-i,-j]$) is mathematically important. In practice, though, you can assume kernels are pre-flipped unless I say otherwise.

Convolving in 2D

Since f and h are defined over finite regions, we can write them out in two-dimensional arrays:

Image (f)

128	54	9	78	100
145	98	240	233	86
89	177	246	228	127
67	90	255	148	95
106	111	128	84	172
221	154	97	69	94

Filter (h)

0.1	0.1	0.1
0.1	0.2	0.1
0.1	0.1	0.1

filter kernel

filter support
footprint

Note: *This is not matrix multiplication*!

The filter values outside the boundary of the filter are always assumed to be zero.

Q: What happens at the boundary of the image?

Normalization

Suppose f is a flat / constant image, with all pixel values equal to some value C.

Image (f)

C	C	C	C	C
C	C	C	C	C
C	C	C	C	C
C	C	C	C	C
C	C	C	C	C
C	C	C	C	C

Filter (h)

h_{13}	h_{23}	h_{33}
h_{12}	h_{22}	h_{32}
h_{11}	h_{21}	h_{31}

normalize $\rightarrow$

$h_{12}/\Sigma h_{ij}$		
$h_{11}/\Sigma h_{ij}$		

Q: What will be the value of each pixel after filtering?

$\sum_{i,j} C h_{ij} = C \cdot \sum h_{ij}$

Q: How do we avoid getting a value brighter or darker than the original image?

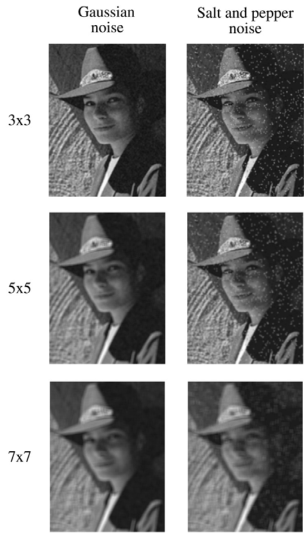
$h \leftarrow \frac{h}{\sum_{ij} h_{ij}}$

Mean filters

How can we represent our noise-reducing averaging as a convolution filter (know as a **mean filter**)?

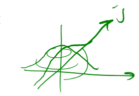
$$\frac{1}{N^2}\underbrace{\left.\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}\right\}N}_{N}$$

Effect of mean filters



Gaussian filters

Gaussian filters weigh pixels based on their distance from the center of the convolution filter. In particular:



$$h[i,j] = \frac{e^{-(i^2+j^2)/(2\sigma^2)}}{C}$$

This does a decent job of blurring noise while preserving features of the image.

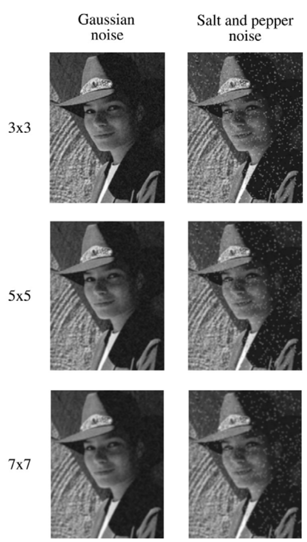
What parameter controls the width of the Gaussian? σ

What happens to the image as the Gaussian filter kernel gets wider? blurrier

What is the constant C? What should we set it to?

$$C = \sum e^{-(i^2+j^2)/(2\sigma^2)}$$

Effect of Gaussian filters

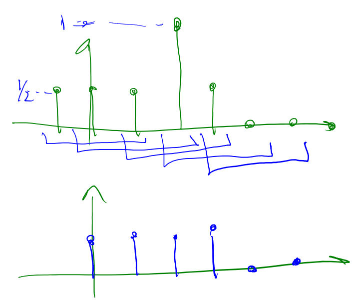


Median filters

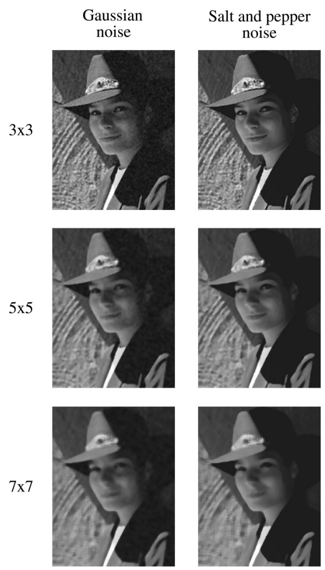
A **median filter** operates over an NxN region by selecting the median intensity in the region.

What advantage does a median filter have over a mean filter? outlier rejection, edge preserving

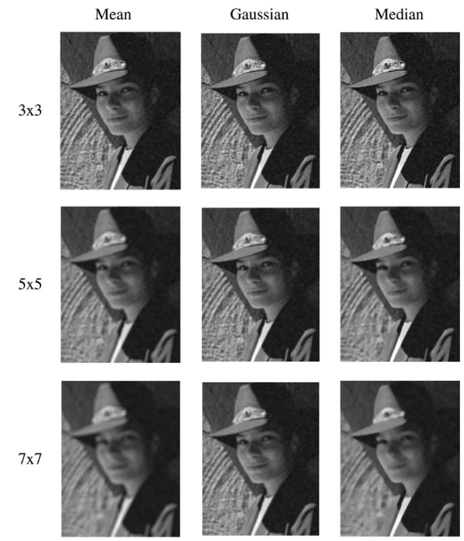
Is a median filter a kind of convolution? N



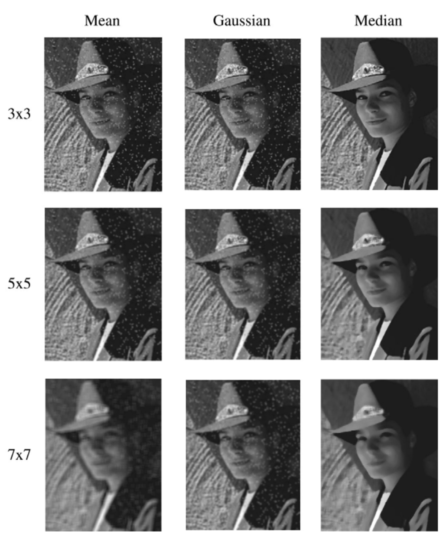
Effect of median filters



Comparison: Gaussian noise

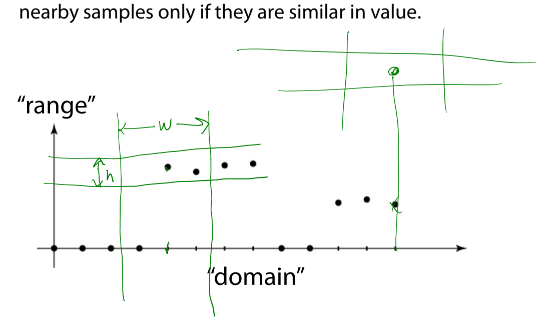


Comparison: salt and pepper noise



Bilateral filtering

Bilateral filtering is a method to average together nearby samples only if they are similar in value.



Q: What happens as the range size becomes large?

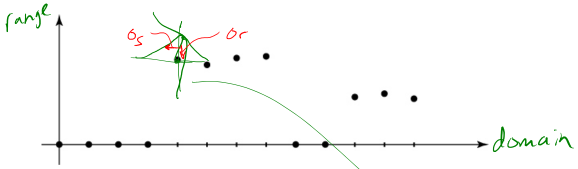
Becomes a mean filter

Q: Will bilateral filtering take care of impulse noise? — or salt + pepper

No

Bilateral filtering

We can also change the filter to something "nicer" like Gaussians:



$$e^{\left(\frac{-i^2}{2\sigma_s^2}+\frac{-r^2}{2\sigma_r^2}\right)} = e^{\frac{-i^2}{2\sigma_s^2}}\cdot e^{\frac{-r^2}{2\sigma_r^2}}$$

Recall that convolution looked like this:

$$g[i] = \sum_k f[k]h[i-k]$$

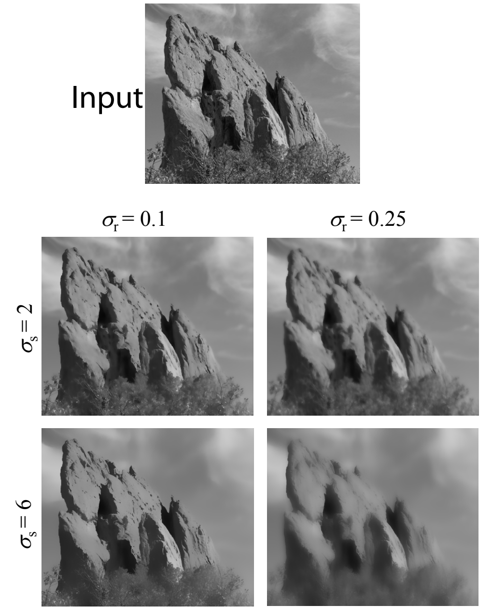
Bilateral filter is similar, but includes both range and domain filtering:

$$g[i] = 1/C\sum_k f[k]h_{\sigma_s}[i-k]\ h_{\sigma_r}(f[i]-f[k])$$

and you have to normalize as you go:

$$C = \sum_k h_{\sigma_s}[i-k]\ h_{\sigma_r}(f[i]-f[k])$$

2D images: $\underbrace{e^{\frac{-(i^2+j^2)}{2\sigma_s^2}}}_{h_{\sigma_s}}\cdot\underbrace{e^{\frac{-r^2}{2\sigma_r^2}}}_{h_{\sigma_r}}$



Paris, et al. SIGGRAPH course notes 2007

Edge detection

One of the most important uses of image processing is **edge detection:**

- Really easy for humans
- Really difficult for computers

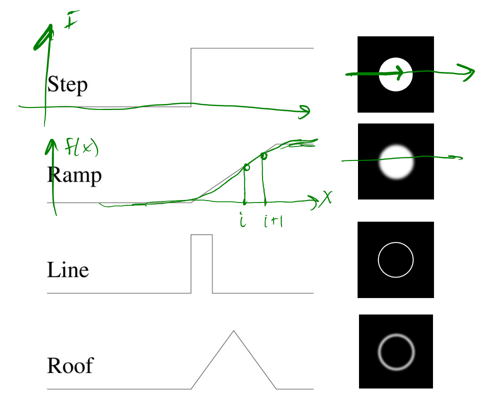
- Fundamental in computer vision
- Important in many graphics applications

What is an edge?

$\tilde{h}_x = [0\ \ -1\ \ 1]$

$h_x = [1\ \ -1\ \ 0]$

$\frac{df}{dx} \approx h_x * f$



finite difference

$$\frac{df}{dx} \approx f[i+1]-f[i] \approx (-1)\cdot f[i] + (1)\cdot f[i+1]$$

Q: How might you detect an edge in 1D?

$\left|\frac{df}{dx}\right| > \text{thresh}$

Gradients

The **gradient** is the 2D equivalent of the derivative:

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

Properties of the gradient

- It's a vector
- Points in the direction of maximum increase of f
- Magnitude is rate of increase

$$\theta = \tan^{-1}\left(\frac{\partial f/\partial y}{\partial f/\partial x}\right)$$

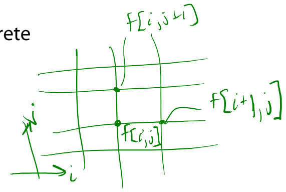
$$\tilde{h}_x = \begin{bmatrix} 0 & -1 & 1 \end{bmatrix} \qquad \tilde{h}_y = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Note: use `atan2(y,x)` to compute the angle of the gradient (or any 2D vector).

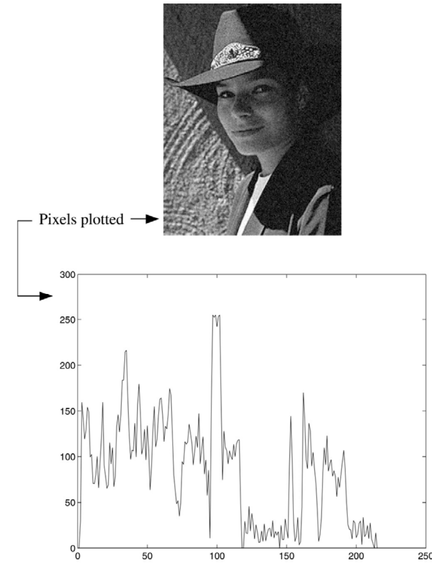
How can we approximate the gradient in a discrete image?

$$\frac{\partial f}{\partial x} \approx f[i+1,j] - f[i,j]$$

$$\frac{\partial f}{\partial y} \approx f[i,j+1] - f[i,j]$$



Less than ideal edges



Steps in edge detection

Edge detection algorithms typically proceed in three or four steps:

- **Filtering**: cut down on noise
- **Enhancement**: amplify the difference between edges and non-edges
- **Detection**: use a threshold operation
- **Localization** (optional): estimate geometry of edges as 1D contours that can pass between pixels

Edge enhancement

central diff.

$$\tilde{h}_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

A popular gradient filter is the **Sobel operator**:

$$\tilde{h}_x = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\tilde{s}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\tilde{s}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\underbrace{h_y * \left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * f\right)}$$

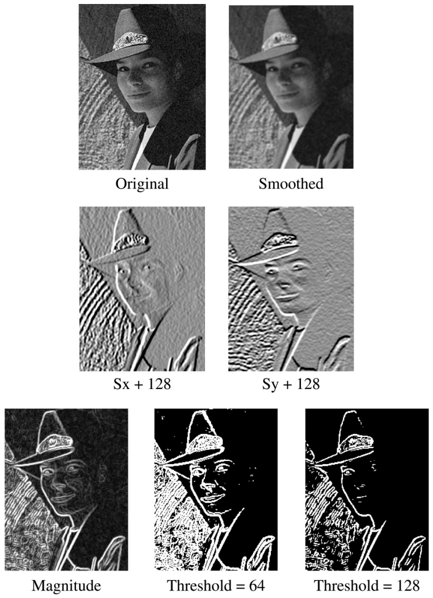
$$\left(\underbrace{h_x * \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}_{S_x}\right) * f \rightarrow S_x * f$$

$$\hat{h}_x \quad \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

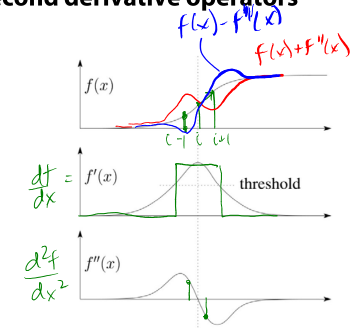
We can then compute the magnitude of the vector $(\tilde{s}_x, \tilde{s}_y)$.

Note that these operators are conveniently "pre-flipped" for convolution, so you can directly slide these across an image without flipping first.

Results of Sobel edge detection



Second derivative operators



$\frac{df}{dx}[i] \approx f[i+1] - f[i]$

$\frac{df}{dx}[i-1] \approx f[i] - f[i-1]$

$\frac{d^2f}{dx^2}[i] \approx f[i-1] - 2f[i] + f[i+1]$

$\tilde{h}_{xx} = [1 \ \ -2 \ \ 1]$

$= h_{xx}$

The Sobel operator can produce thick edges. Ideally, we're looking for infinitely thin boundaries.

An alternative approach is to look for local extrema in the first derivative: places where the change in the gradient is highest.

Q: A peak in the first derivative corresponds to what in the second derivative? 0

Q: How might we write this as a convolution filter?

Constructing a second derivative filter

We can construct a second derivative filter from the first derivative.

First, one can show that convolution has some convenient properties. Given functions *a, b, c*:

Commutative: $a * b = b * a$

Associative: $(a * b) * c = a * (b * c)$

Distributive: $a * (b + c) = a * b + a * c$

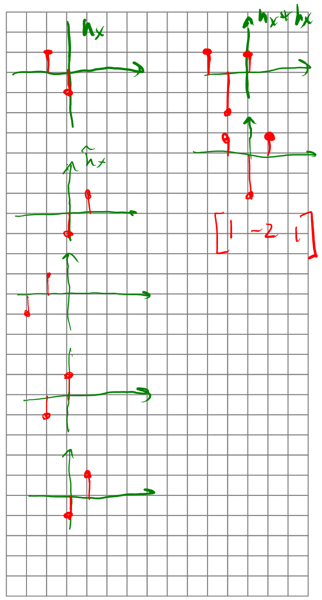
The "flipping" of the kernel is needed for associativity. Now let's use associativity to construct our second derivative filter…

$h_x = [1 \ \ -1 \ \ 0]$

$\frac{d^2f}{dx^2} = \frac{d}{dx}\left(\frac{d}{dx} f\right)$ $\quad \tilde{h}_x = [0 \ \ -1 \ \ 1]$

$\approx h_x * (h_x * f)$

$(h_x * h_x) * f$



Localization with the Laplacian

An equivalent measure of the second derivative in 2D is the **Laplacian**:

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \approx h_{xx} * f + h_{yy} * f$$

$(h_{xx} + h_{yy}) * f$

Using the same arguments we used to compute the gradient filters, we can derive a Laplacian filter to be:

$[1 \ \ -2 \ \ 1] + \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

$$\Delta = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

(The symbol Δ is often used to refer to the *discrete* Laplacian filter.)

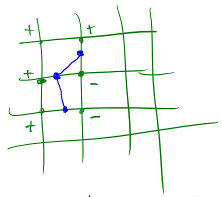
Zero crossings in a Laplacian filtered image can be used to localize edges.

Localization with the Laplacian

Original

Smoothed

Laplacian (+128)



"marching squares"

Sharpening as blur removal

We can also think of sharpening as blur removal:

$$g = K\left(f - \alpha(b * f)\right)$$

Suppose:

$$\alpha = 5/6 \qquad b = \begin{bmatrix} 0 & 1/5 & 0 \\ 1/5 & 1/5 & 1/5 \\ 0 & 1/5 & 0 \end{bmatrix}$$

We can let α depend on ∇f to sharpen selectively, a.k.a., **unsharp masking**. How should α vary with ∇f ?