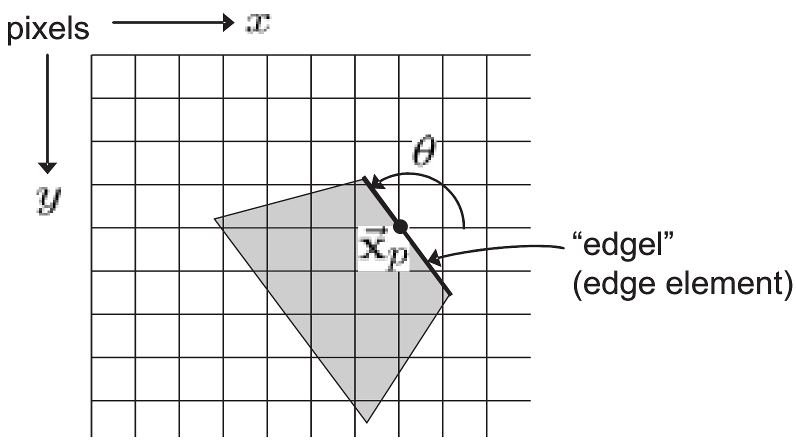
Edge Detection

Goal: Detection and Localization of Image Edges.

Motivation:

- Significant, often sharp, contrast variations in images caused by illumination, surface markings (albedo), and surface boundaries. These are useful for scene interpretation.
- **Edgels (edge elements):** significant local variations in image brightness, characterized by the position $\vec{\mathbf{x}}_p$ and the orientation θ of the brightness variation. (Usually $\theta \mod \pi$ is sufficient.)



- **Edges:** sequence of edgels forming smooth curves

Two Problems:

1. estimating edgels
2. grouping edgels into edges

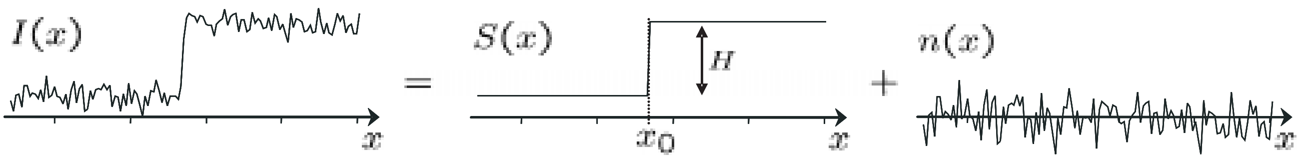
Matlab Tutorials: cannyTutorial.m

1D Ideal Step Edges

Assume an ideal step edge corrupted by additive Gaussian noise:

$$I(x) = S(x) + n(x) \ .$$

Let the signal S have a step edge of height H at location x_0, and let the noise at each pixel be Gaussian, independent and identically distributed (IID).



Gaussian IID Noise:

$$n(x) \sim N(0, \sigma_n^2) \ , \qquad p_n(n; 0, \sigma_n^2) \ = \ \frac{1}{\sqrt{2\pi}\sigma_n} \, e^{-n^2/\sigma_n^2}$$

Expectation:

$$\text{mean: } \mathrm{E}[n] \ \equiv \ \int n \, p_n(n) \, dn = 0$$

$$\text{variance: } \mathrm{E}[n^2] \ \equiv \ \int n^2 \, p_n(n) \, dn = \sigma_n^2$$

Independence: $p(n(x_1), n(x_2)) = p(n(x_1))p(n(x_2))$ for $x_1 \neq x_2$.

$$\mathrm{E}[n(x_1)\, n(x_2)] = \sigma_n^2 \delta_{x_1,x_2} = \begin{cases} 0 & \text{when } x_1 \neq x_2 \\ \sigma_n^2 & x_1 = x_2 \end{cases} \qquad (1)$$

Remark: Violations of the main assumptions, i.e., the idealized step edge and additive Gaussian noise, are commonplace.

Signal Response

Assume a linear filter, with impulse response $f(x)$:

$$\begin{aligned} r(x) \;=\; f(x) * I(x) \;&=\; f(x) * S(x) \;+\; f(x) * n(x) \\ &=\; \quad r_S(x) \quad + \quad r_n(x) \end{aligned}$$

So the response is the sum of responses to the signal and the noise.

We will use large values of the absolute response $|r(x)|$ to detect edges.

Therefore we want $|r(x)|$ to be small when $H = 0$ (i.e. no step, $S(x) = c$ for a constant c). So we require

$$f * c = 0$$

for any constant c. Equivalently,

$$\sum_k f(k) = 0. \tag{2}$$

Thus the filter kernel has zero DC response ("DC" denotes a signal with frequency 0).

Response to Noise

The mean and variance of the response to noise $r_n(x)$,

$$r_n(x) \;=\; \sum_{k=-K}^{K} f(-k)\, n(x+k)\;,$$

where K is the radius of filter support, can be shown to be (see notes on next page)

$$\begin{aligned}
\mathrm{E}[r_n(x)] \;&=\; \sum_k f(-k)\, \mathrm{E}[n(x+k)] \;=\; 0 \\
\mathrm{E}[r_n^2(x)] \;&=\; \sum_k \sum_l f(-l)\, f(-k)\, \mathrm{E}[n(x{+}k)n(x{+}l)] \\
&=\; \sigma_n^2 \sum_k f^2(k)
\end{aligned}$$

Note that the standard deviation of the noise response,

$$\left(\mathrm{E}[r_n^2(x)]\right)^{1/2} = \sigma_n \sqrt{\sum_k f^2(k)},$$

depends only on the 2-norm of the filter kernel $f(k)$, not on the detailed shape of the kernel, nor on the pixel x.

Expectation of Sums and Products of Random Variables

Suppose n_1 and n_2 are two random variables with the joint probability distribution $p(n_1, n_2)$. If the variables are independent then this joint distribution can be written as the product of the individual probability distributions $p(n_1)$ and $p(n_2)$, namely $p(n_1, n_2) = p(n_1)p(n_2)$. In anycase, we have the general marginalization property of probability distributions

$$p(n_1) = \int p(n_1, n_2) dn_2, \quad \text{and } p(n_2) = \int p(n_1, n_2) dn_1. \tag{3}$$

Note this is easy to show for independent random variables n_1 and n_2.

Let a, b be two constants. Then it follows that $\mathrm{E}[an_1 + bn_2] = a\mathrm{E}[n_1] + b\mathrm{E}[n_2]$. In particular, the expectation of a sum of random variables is just the sum of the expectation of each term. The variables n_1 and n_2 don't need to be independent. The derivation of this is as follows,

$$\begin{aligned}
\mathrm{E}[an_1 + bn_2] &\equiv \int\int (an_1 + bn_2) p(n_1, n_2) dn_1 dn_2, \\
&= \int\int an_1 p(n_1, n_2) dn_1 dn_2 + \int\int bn_2 p(n_1, n_2) dn_1 dn_2, \\
&= \int an_1 p(n_1) dn_1 + \int bn_2 p(n_2) dn_2, \text{ by marginalization}, \\
&= a \int n_1 p(n_1) dn_1 + b \int n_2 p(n_2) dn_2, \\
&= a\mathrm{E}[n_1] + b\mathrm{E}[n_2].
\end{aligned} \tag{4}$$

Note that here we just used the marginalization property of $p(n_1, n_2)$, and not independence of n_1 and n_2.

In contrast, it is **not generally** the case that the expectation of products of random variables is just the product of the expectations. For example, $\mathrm{E}[n_1 n_1] = \sigma_n^2 \neq \mathrm{E}[n_1]\mathrm{E}[n_1] = 0$. However, for **independent** random variables n_1 and n_2 we have

$$\begin{aligned}
\mathrm{E}[n_1 n_2] &\equiv \int\int n_1 n_2 p(n_1, n_2) dn_1 dn_2, \\
&= \int\int n_1 n_2 p(n_1) p(n_2) dn_1 dn_2, \quad \text{by independence}, \\
&= \left(\int n_1 p(n_1) dn_1\right) \left(\int n_2 p(n_2) dn_2\right), \\
&= \mathrm{E}[n_1]\mathrm{E}[n_2] \text{ for independent } n_1, n_2.
\end{aligned} \tag{5}$$

Expectation and Variance of Noise Response $r_n(x)$

Using equation (4) (and its extension to sums of more than two random variables), we find

$$E[r_n(x)] = E[\sum_{k=-K}^{K} f(-k)n(x+k)] = \sum_{k=-K}^{K} f(-k)E[n(x+k)] = 0.$$

Similarly,

$$\begin{aligned}
E[r_n^2(x)] &= E[\{\sum_{k=-K}^{K} f(-k)n(x+k)\}\{\sum_{j=-K}^{K} f(-j)n(x+j)\}], \\
&= E[\sum_{k=-K}^{K}\sum_{j=-K}^{K} f(-k)n(x+k)f(-j)n(x+j)], \\
&= \sum_{k=-K}^{K}\sum_{j=-K}^{K} f(-k)f(-j)E[n(x+k)n(x+j)], \\
&= \sum_{k=-K}^{K}\sum_{j=-K}^{K} f(-k)f(-j)\sigma_n^2\delta_{k,j}, \text{ by equation (1)}, \\
&= \sum_{k=-K}^{K} f^2(-k)\sigma_n^2,
\end{aligned}$$

which is the desired result reported on p.4.

Signal to Noise Ratio

What is the optimal linear filter for the detection and localization of a step edge in an image?

We might measure how well we are doing by comparing $|\mathrm{E}[r(x_0)]|$ (i.e., the absolute value of expected value of the signal at a step edge x_0), to the standard deviation of the noise in this response, $\sqrt{\mathrm{E}[r_n^2(x_0)]}$.

From the analysis above, we have

$$|\mathrm{E}[r(x_0]| = |r_s(x_0) + \mathrm{E}[r_n(x_0)]| = |r_s(x_0)| = |f * S(x_0)|.$$

And

$$\mathrm{E}[r_n^2(x_0)] = \sigma_n^2 \sum_{k=-K}^{K} f^2(k).$$

So we define *Signal-to-Noise Ratio* (SNR) to be:

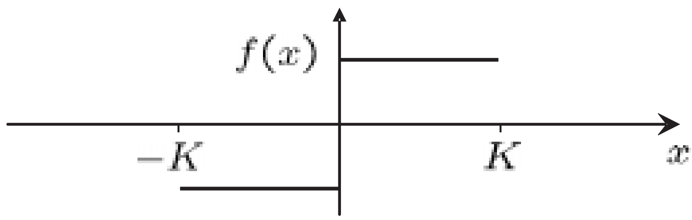
$$SNR \; = \; \frac{|(f * S)(x_0)|}{\sigma_n \sqrt{\sum_k f^2(k)}}$$

Note the SNR is invariant to scaling f. That is, replacing $f(k)$ by the filter $af(k)$ gives the same SNR for any constant $a \neq 0$.

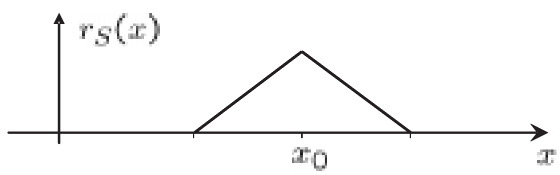
Criteria for Optimal Filters

Criterion 1: *Good Detection.* Choose the filter to maximize the SNR of the response at the edge location, subject to constraint that the responses to constant signals are zero.

For a filter with a support radius of K pixels, the optimal filter is a matched filter, i.e., a difference of square box functions:



Response to ideal step:



Explanation:
Assume, with out loss of generality that $\sum f^2(x) = 1$, and to ensure zero DC response, $\sum f(x) = 0$.
Then, to maximize the SNR, we simply maximize the inner product of $S(x)$ and the impulse response, reflected and centered at the step edge location, i.e., $f(x_0 - x)$.

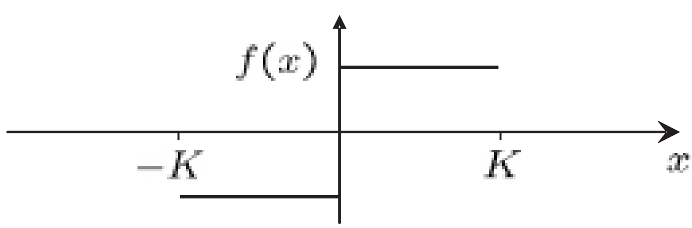
Criteria for Optimal Filters (cont)

Criterion 2: *Good Localization.* Let $\{x_l^*\}_{l=1}^{L}$ be the local maxima in response magnitude $|r(x)|$. Choose the filter to minimize the root mean squared error between the *true edge location* and the *closest peak* in $|r|$; i.e., minimize

$$LOC = \frac{1}{\sqrt{\mathrm{E}[\min_k |x_l^* - x_0|^2]}}$$

Caveat: for an optimal filter this does not mean that the closest peak should be the most significant peak, or even readily identifiable.

Result: Maximizing the product, $SNR \cdot LOC$, over all filters with support radius K produces the same matched filter already found by maximizing SNR alone.



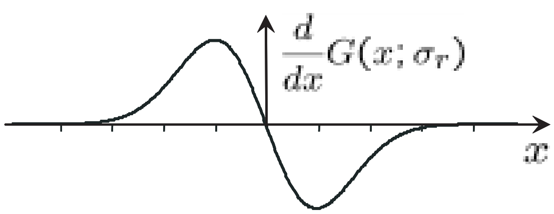
Criteria for Optimal Filters (cont)

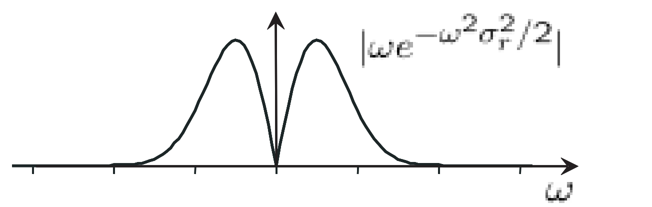
Criterion 3: *Sparse Peaks.* Maximize $SNR \cdot LOC$, subject to the constraint that peaks in $|r(x)|$ be as far apart, on average, as a manually selected constant, $xPeak$:

$$\mathrm{E}[\,|x^*_{k+1} - x^*_k|\,] \;=\; xPeak$$

When $xPeak$ is small, $f(x)$ is similar to the matched filter above. But for $xPeak$ larger (e.g., $xPeak \approx K/2$) then the optimal filter is well approximated by a derivative of a Gaussian:

$$f(x) \;\approx\; \frac{dG(x;\sigma_r)}{dx} \;=\; \frac{-x}{\sqrt{2\pi}\sigma_r^3}\, e^{-\frac{x^2}{2\sigma_r^2}}\,, \text{ with } \mathcal{F}\left[\frac{dG(x;\sigma_r)}{dx}\right] \;=\; i\,\omega\, e^{-\frac{\omega^2\sigma_r^2}{2}}$$





Conclusion:

Sparsity of edge detector responses is a critical design criteria, encouraging a smooth envelope, and thereby less power at high frequencies. The lower the frequency of the pass-band, the sparser the response peaks.

There is a one parameter family of optimal filters, varying in the width of filter support, σ_r. Detection (SNR) improves and localization (LOC) degrades as σ_r increases.

Multiscale Edge Features

Multiple scales are also important to consider because salient edges occur at multiple scales:

1) Objects and their parts occur at multiple scales:

2) Cast shadows cause edges to occur at many scales:

3) Objects may project into the image at different scales: