Solutions to Homework #2

Q5.1. In template matching, a template has to be explicitly compared to an image. Unfortunately, a template can be in any possible orientation and scale. Consequently, rotated, translated and scaled (and maybe warped) versions of the template have to be compared (cross-correlated) to an image in order to evaluate the goodness of the match. This is not required in a Hough transform which is a voting procedure. To some extent though, this comparison is between apples and oranges. A Hough transfom approach works well when there is a parametric form of the shape (line, ellipse etc.) and therefore the number of parameters in the voting space is not large. This is true in the case of lines which require only two parameters to fully specify them. If the template is not an image, but is instead the 2D locations of points, then template matching can be converted into a Hough transform like approach by asking each data edge point to vote when it "hits" or comes close to hitting a template point. So, even though template matching with its roots in cross-correlation and the Hough transform with its roots in voting seem irreconcilable at first glance, they start becoming similar when the template is not an image but is a *configuration* of points or other geometric primitives.

Q5.5. The short answer to this question is that the $(\rho\,\theta)$ parameterization keeps both quantities finite as opposed to the (m, n) representation. This allows the finite space to be better discretized (since one cannot evenly sample an infinite interval). Assume that you are restricted to a total of L^2 bins in the Hough space and you are asked to keep the number of bins the same in both. A finite interval $\rho \in [0, p_{max}]$, will have a bin width of ρ_{max}/L and a finite interval $\theta \in [0, \pi]$ will have a bin width π/L. In contrast, one has to accept an *a priori* tradeoff in discretizing m and n. If M and N are chosen as limits, then the bin widths are M/L and N/L which will usually be much larger than the corresponding bin widths in the $(\rho\,\theta)$ case. And, we run the risk of not covering lines corresponding to large values of m and n. The bin width difference is then a characterization of performance with the additional caveat that some lines detected by the finite discreization will be missed by the infinite discretization.

Q5.10. If the point to be avoided is (x, y), then a reasonable choice for the energy term is $\int_0^S \frac{1}{[(x(s)-x)^2+(y(s)-y)^2]} ds$. If the snake attempts to come close to the point (x, y), then the energy of the snake shoots up to infinity. Since we are minimizing the snake energy, the snake will avoid coming close to (x, y).

E5.1. This was discussed rather extensively in class. The main idea is that there is additional information at each edge point (x_i, y_i) in the form of the gradient direction θ_i. Consequently, a line is deemed to be more likely to pass through point (x_i, y_i) if its slope is close to θ_i. We require a measure of this closeness. One possibility is $\cos^2(\theta_i - \theta)$ where both θ_i and θ are in the interval $[0, \pi)$. The use of a trigonometric measure allows us to treat equally, θ_i and $\pi - \theta_i$. Given this measure of closeness, we obtain a weighted (real numbered) vote in Hough space.

E5.2. The equation of the curve in Hough space $(\rho\,\theta)$ is $\rho = x\cos\theta + y\sin\theta$. This equation still has to respect limits $\theta \in [0, \pi)$ and $\rho \in [0, \rho_{max}]$. A better way of expressing this equation is

$$\frac{\rho}{\rho_{max}} = \frac{x}{\rho_{max}}\cos\theta + \frac{y}{\rho_{max}}\sin\theta = \cos[(\phi + \theta)/2]$$

where $\phi \stackrel{def}{=} \tan^{-1}(y/x)$. This is by definition in the first quadrant since x and y are always positive. Since ρ is non-negative, a constraint on $\phi + \theta$ is that it lies in the interval $[-\pi, \pi]$ or otherwise $\cos(\phi + \theta) < 0$. This constrains θ to lie in the interval $[-\pi - \phi\, \pi - \phi]$. However, the other constraints on θ also have to be respected.

Extra question. $E(f) = \int_0^1 (f')^2 dx$. The Euler-Lagrange equation for this is $\frac{d}{dx}\frac{\partial L}{\partial f_x} = 0$ which yields $\frac{\partial^2 f}{\partial x^2} = f'' = 0$. An obvious discretization of f'' is $f_{i+1} - 2f_i + f_{i-1}$ which can be obtained by applying differencing to $f_{i+1} - f_i$. Now, we may discretize the functional itself by using forward differencing $E_{disc}(f) = \sum_i (f_{i+1} - f_i)^2$. When we collect terms w.r.t. f_i, we get two terms $(f_{i+1} - f_i)^2$ and $(f_i - f_{i-1})^2$. When we differentiate terms only depending on f_i and set the result to zero, we get $f_i - f_{i+1} + f_i - f_{i-1} = 2f_i - f_{i+1} - f_{i-1} = 0$ which is the same as obtained by discretizing the Euler-Lagrange equation.

E5.8. This is a very poorly worded question and is correct only for curves which have the following property $x(s) = s + b$. Then the curvature $\kappa(s) = |x'y'' - y'x''| = |y''|$ and when discretized and squared, the energy term is $|p_{i+1} - 2p_i + p_{i-1}|^2$. Note that p_i consists of the pair x_i, y_i and that x is evenly spaced which means its second derivative is zero.