Midterm CMSC 828
Instructor: David Jacobs
Distributed 10/21/04, due 10/28/04

1. Fourier Transforms and Smoothing
    a. Let *f* be a periodic function on the interval $-\pi$ to π so that $f(t) = 1$ for $t <= 0$ and $f(t) = 2$ for $t > 0$. Express *f* as a Fourier series. That is, give a0, a1, …, b1, b2, … so that $f(t) = a0/2 + a1cos(t) + b1sin(t) + a2 cos(2t) + b2sin(2t) + ...$
    b. Suppose that $n(t)$ represents some noise added to this signal. We'll pick a very simple example of noise. Suppose $n(t) = k$ for $1 <= t <= 1 + \varepsilon$ and $n(t) = 0$ otherwise. *k* is a random variable drawn from a uniform distribution between -.5 and .5. Express $n(t)$ as a Fourier series.
    c. $G*f$ denotes the convolution of *f* with a Gaussian. $G*n$ is similarly defined. Show that $||G*f||/||G*(f+n)|| > ||f||/||f+n||$. If this isn't always true, indicate what additional assumptions are needed to make it true.
    d. Consider the following simple segmentation algorithm. All pixels with a value above 1.5 are assigned to one segment, and the other pixels are assigned to a second segment. Suppose we apply this to a 1D image consisting of $f(t)$ with independent, identically distributed Gaussian noise added to each pixel. Explain why and under what circumstances we might improve the performance of this algorithm by first smoothing the image with a Gaussian. A qualitative answer may be necessary for this question, but the more precise you can be, the better.
2. Diffusion. Let $f(t,0)$ be an initial distribution of temperature. Suppose $f(t,0)$ has exactly two local maxima. The temperature evolves over time according to a homogenous, isotropic diffusion.
    a. At some subsequent time, can *f* have more than two local maxima? Either prove that it cannot or give an example demonstrating that it can.
    b. **Challenge problem:** For extra credit, answer question (a) for Perona-Malik diffusion.
3. Markov Processes. Suppose Alice and Bob are playing a game. There is a token that starts with Alice. Alice and Bob each have a coin. Whoever has the token, flips their coin. If the coin lands heads, they pass the token to the other player. If the coin lands tails, the player keeps the token. This is repeated one million times. Whoever has the token at the end wins. You get to watch the first move of a few thousand games, and conclude that Alice is using a biased coin, that lands tails 60% of the time.
    a. If Bob's coin is fair, what is the probability that Alice will win the game?
    b. Suppose you also get to learn the outcome of a few thousand games, and discover that Bob is winning 60% of the time. Use this information to determine the probability that Bob's coin lands tails when it is flipped.
4. Belief Propagation.
    a. Consider the game described in Problem 3. Suppose Alice and Bob are both using coins that land tails 60% of the time. Suppose further that at

the end of the third turn, Bob has the token. What is the probability that Bob had the token at the end of the second turn?

b. Suppose we have a 1D image that is generated in the following way. We first produce a signal, and then add noise to it. To get the signal, the first pixel has an integer value drawn from a uniform distribution between 0 and 255. Each subsequent pixel, with probability 49/50, has the same value as the previous pixel. With probability 1/50, a new pixel has a value drawn from a uniform distribution from 0 to 255. Next we add noise to the signal. For each pixel, we draw the noise from an i.i.d. Gaussian distribution. We would like to take the noisy signal, and segment the pixels into groups that originally had the same intensity, before noise was added. Explain how to do this using belief propagation. Be explicit about what the nodes in the belief net will be, what the causal links will be, and what the conditional probabilities will be.

c. **Challenge problem:** For extra credit, implement belief propagation using the belief net that you describe above. Also implement Perona-Malik diffusion, and describe a method of using it for segmentation on images like this. Compare the performance of the two approaches on synthetic data. Partial credit will be given for doing part of this.