Statistical and Learning Techniques in Computer Vision
Homework 3: Due Thursday September 21, 2006

1. (**10 points**) Prove that when p_A and p_B are independent, $H(A, B) = H(A) + H(B)$, whereas when p_A and p_B are perfectly correlated $H(A, B) = H(A) = H(B)$.

   Note that the correlation coefficient between two random variables, $\mathbf{x}$ and $\mathbf{y}$ is

   $$\mathrm{corr}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{E}[(\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y)]}{\sigma_x \sigma_y},$$

   where $\mu_x = \mathbf{E}[\mathbf{x}]$, $\mu_y = \mathbf{E}[\mathbf{y}]$, $\sigma_x^2 = \mathbf{E}[(\mathbf{x} - \mu_x)^2]$ and $\sigma_y^2 = \mathbf{E}[(\mathbf{y} - \mu_y)^2]$. Two images are perfectly correlated if their corresponding pixels (pixels at the same location in each image) each have a correlation coefficient of 1.0.

2. (**15 points**) Consider the two example "images" shown in Figure 1. Based on the Mutual Information image comparison measure defined in Lecture 3, what is the optimal rotation and translation between the two image? Justify your answer as carefully as you can. How does your answer change when the image comparison is the sum of the squared difference in intensities between the two images?

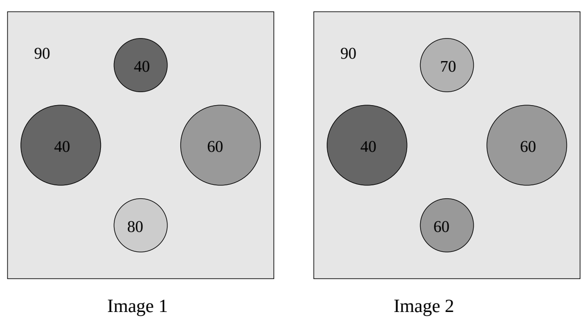

Figure 1: Two images with the same geometric shapes but different intensities. The numbers shown in each image region indicate the intensity of each pixel in that region. Assuming the center of the image coordinate system is in the geometric center of the two images (as opposed to the upper left corner), the four circles are at locations $(c, 0)$, $(0, c)$, $(-c, 0)$ and $(0, -c)$. The left and right circles both have radius r_h, and the top and bottom circles both have radius r_v, with $r_h > r_v$.

3. (**10 points**) A Markov Chain is a sequence of random variables $\mathbf{x}_1, \mathbf{x}_2.\mathbf{x}_3 \ldots$ such that the present state $\mathbf{x}_t$ only depends on the previous state, i.e.

   $$p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{x}_{t-1}, \ldots, \mathbf{x}_0) = p(\mathbf{x}_{t+1}|\mathbf{x}_t) \ .$$

   As you can see, a Markov Chain is a special case of a Markov Random Field. If the state space of the random variables $\mathbf{x}_t$ is finite then the conditional probabilities $p(\mathbf{x}_{t+1}|\mathbf{x}_t)$ are represented by a matrix, called the transition matrix P. The element (i, j) of the matrix P is equal to

$$p_{ij} = \text{prob}(\mathbf{x}_{t+1} = j | \mathbf{x}_t = i) \ .$$

(a) Represent the Markov Chain as a graph.

(b) Show that

$$P^{t+s}_{i,j} = \sum_k P^t_{i,k} P^s_{k,j} \ ,$$

where $P^c_{i,j} := \text{prob}(\mathbf{x}_{t+c} = j | \mathbf{x}_{t+c} = i)$. Tip: Prove the above equation for the case $s = 2$. Then use induction to show the more general statement.

4. (**20 points - Markov Chain example**) Tyler, Robert, and Jane meet every day of the year to have lunch at a nearby restaurant. There are 8 restaurants in the area they like to go to. Jane proposed a random process so that they can simplify the way they choose their restaurant on any particular day. Based on historical data she uses the following transition matrix to design the process:

$$P = \begin{pmatrix} 0.3 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.05 & 0.05 & 0.4 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.2 & 0.4 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.1 & 0.1 & 0.2 & 0.2 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.3 & 0.1 \\ 0.2 & 0.1 & 0.2 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.05 & 0.05 & 0.05 & 0.2 & 0.2 & 0.05 & 0.2 & 0.2 \end{pmatrix}$$

(a) Asssuming that each of the restaurants is chosen with equal probability write a simulator so that you can generate a sequence of 360 days.

(b) Compute the autocovariance of two different random variables $\mathbf{x}_t$ and $\mathbf{x}_{t+c}$ using the sample data you generated. The autocovariance is

$$R(c) = \frac{1}{(T-c)\sigma^2} \sum_{t=1}^{T-c} (\mathbf{x}_t - \mu)(\mathbf{x}_{t+c} - \mu) \ ,$$

where T is the total length of your sample sequence, i.e. 360. Plot the autocorrelation $R(\cdot)$ for different values of $c = 1, 2, 3, \ldots$. Describe what you observe.