# CS-543: Computer Vision Instructor: D.A. Forsyth

**Homework 4**

## Instructions

This homework is a check, and should be done individually. It is due in two weeks from handout, i.e. 28 April 2009. Submit by emailing a PDF to Alex.

## Question 1: Tracking

We have a linear dynamical system with states $\mathbf{X}_i$. In particular, $\mathbf{X}_i$ is a normal random variable with mean $\mathcal{D}\mathbf{X}_{i-1}$ and covariance $\Sigma_d$. At each time step, there are *two* components to the measurement $\mathbf{Y}_i$. The components are $\mathbf{Y}_{i1}$ and $\mathbf{Y}_{i2}$. One of these two components is a normal random variable with mean $\mathcal{M}\mathbf{X}_i$ and covariance $\Sigma_m$. The other is a normal random variable with mean $0$ and covariance $\Sigma_n$. We do not know which of the components is informative. $P(\mathbf{X}_1)$ is Gaussian.

**Part 1:** Show that, if $P(\mathbf{X}_{i-1}|\mathbf{Y}_1, ..., \mathbf{Y}_{i-1})$ is a mixture of $k$ Gaussians, then $P(\mathbf{X}_i|\mathbf{Y}_1, ..., \mathbf{Y}_{i-1})$ is a mixture of $k$ Gaussians.

**Part 2:** Show that, if $P(\mathbf{X}_i|\mathbf{Y}_1, ..., \mathbf{Y}_i)$ is a mixture of $k$ Gaussians, then

$$P(\mathbf{X}_i|\mathbf{Y}_1, ..., \mathbf{Y}_i, \text{first component of measurement is informative})$$

is a mixture of $k$ Gaussians.

**Part 3:** Show that the posterior, $P(\mathbf{X}_i|\mathbf{Y}_1, ..., \mathbf{Y}_i)$ is a mixture of Gaussians, with $2^i$ mixture components.

**Part 4:** We cannot use an exact representation of the posterior in this case, because there are two many components. Suggest a strategy for managing this problem.

## Question 2: Finding Pedestrians

Read the paper "Histograms of Oriented Gradients for Human Detection" by Navneet Dalal and Bill Triggs (International Conference on Computer Vision & Pattern Recognition - June 2005), and prepare a brief (one page) evaluation and critique. What are the major strengths of the method? what are its major weaknesses? what should be done about them?

## Question 3: Pictorial structures

In class, I described a method called pictorial structures, which is used to find the configuration of a person in an image given a model of appearance and of inter segment relations.

**Part 1:** What problems would result using this method for a lateral view of a person in a walking configuration *where you can see both legs*? (hint: one leg is usually better than the other). What can be done about this?

**Part 2:** Pictorial structure models preserve a tree of spatial relations between body components. Is this tree a sufficient representation? what other relations might exist between body parts? what might be the consequences of ignoring them?