EEL806: Minor II

October 6, 2013

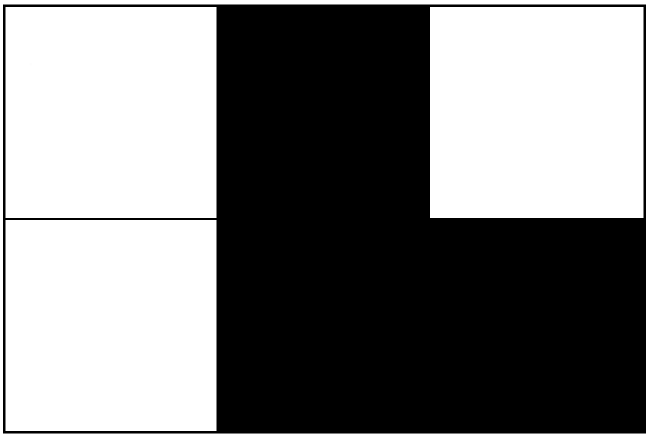
Maximum Marks: 20

1. Describe how a 3-D model of a building can be created from its photograph as shown below. **[5]**

2. (a) What is the key challenge in recognition? Explain with examples. **[2]**

   (b) Why are local features useful for object recognition? What advantages do they have over global features? **[2]**

3. Consider an image which contains only two types of pixels, black and white, *e.g.* as below.



Supposing I define a generative model for such images as follows. There are K world states, such that each pixel comes from one of them. Corresponding to the k^{th} world state there is just one parameter: μ_k, the probability of that state generating a white pixel. The proabability of a black pixel from that state is thus given by $1 - \mu_k$ (this is known as the *Bernoulli distribution* with parameter μ_k). Furthermore, I assume there is a prior probability for each world state, denoted π_k, such that $\pi_k > 0$ and $\sum_{k=1}^{K} \pi_k = 1$.

Let $x \in \{0, 1\}$ denotes the colour of any given pixel, where 0 represents black and 1 represents white. Then, if I know that this pixel comes from the k^{th} world state, I can write the probability distribution over its colour as follows (note that this is just a compact way of writing the Bernoulli distribution as defined above):

$$p(x|worldstate = k) = \mu_k^x (1 - \mu_k)^{1-x}.$$

(a) Given that the world state is unknown, what is $p(x)$, the probability distribution for x under this model? What kind of model is this? **[2]**

(b) Introduce a latent/hidden variable $\mathbf{w}$ for this model, to denote the world state of any given pixel: $\mathbf{w}$ is a binary vector of length K, with only a single entry set to 1 to denote a particular world state, and the remaining entries set to 0. Write down $p(x, \mathbf{w})$, the joint distribution of data and world state. **[1]**

(c) Now suppose we have observed an image with N pixels. Let us denote the entire image data by $\mathbf{X} = \{x_1, x_2, ..., x_N\}$, where x_n denotes the colour of the n^{th} pixel. Also let $\mathbf{W} = \{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_N\}$ denote correspondingly the world states for all pixels. What is $\log p(\mathbf{X}, \mathbf{W})$, *i.e.*, the *complete-data log likelihood*? **[1]**

(d) In order to simultaneously estimate both the model parameters $\{\mu_k\}$ and $\{\pi_k\}$ as well as the hidden world states $\{\mathbf{w}_n\}$, we need to use the EM algorithm. In the E-step, we assume the parameter values are given to us, and compute the posteriors over the world states, denoted $\gamma(w_{nk}) = p(w_{nk} = 1|x_n)$. Use Bayes' theorem to obtain an expression for $\gamma(w_{nk})$. **[2]**

(e) In the M-step, we use the computed posteriors from the E-step to re-estimate the parameter values that maximise the log likelihood of the data. For this model, the M-step updates turn out to be:

$$\mu_k = \frac{\sum_{n=1}^{N} \gamma(w_{nk}) x_n}{\sum_{n=1}^{N} \gamma(w_{nk})};$$

$$\pi_k = \frac{\sum_{n=1}^{N} \gamma(w_{nk})}{N}.$$

Interpret each of these in words. **[2]**

4. Which one of the following cars do you think can be more accurately tracked using Feature Points? Why? **[3]**

(a)

(b)