UNIVERSITY COLLEGE LONDON

University of London

EXAMINATION FOR INTERNAL STUDENTS

2000-2001

For the following qualifications :-

M.Res

COURSE CODE	:	**PPP**
TITLE OF EXAMINATION	:	**Physics, Psychophysics and Physiology of Vision**
DATE	:	**8 March-2001**
TIME	:	**14.00**
TIME ALLOWED	:	**2 hours 30 minutes**

*Answer **three** questions in total. Each question is worth 33 marks. The total time allowed is two and a half hours.*

Question 1

(a)

(i) Write down how the colour signal $\mathbf{C}(\mathbf{x}) = \left(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x})\right)$ at pixel $\mathbf{x}$ in each of the channels of an RGB colour camera is related to the light flux incident on the sensor surface or image irradiance, $E(\mathbf{x}, \lambda)$. Explain the meaning of λ and of any terms you introduce.

(ii) Explain why the total intensity I may be represented as the sum, $R+G+B$ and explain under what conditions it represents the total energy flux in the incident radiation ?

(iii) By considering how the signal in each colour channel is changed when the spectral distribution of the image irradiance changes, show that the effect on $\mathbf{C}(\mathbf{x})$ may be represented as a matrix-vector multiplication,

$$\mathbf{C}'(\mathbf{x}) = \left(R'(\mathbf{x}), G'(\mathbf{x}), B'(\mathbf{x})\right) = \mathbf{C}(\mathbf{x})\mathbf{A}(\mathbf{x}) = \left(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x})\right)\mathbf{A}(\mathbf{x}) \tag{1}$$

where $\mathbf{C}(\mathbf{x})$ is regarded as a row-vector and the matrix $\mathbf{A}(\mathbf{x})$ is diagonal. Describe what type of transformation of $\mathbf{C}(\mathbf{x})$ this matrix multiplication represents in the RGB colour space.

(iv) Describe why, in spite of the result in (iii) above, the colour signal $\mathbf{C}(\mathbf{x})$ may not properly be regarded as a vector and show that the form of the diagonal matrix $\mathbf{A}$ is further constrained in a manner consistent with this fact.

[10 marks]

[Question 1 cont. over page]

[TURN OVER]

[Question 1 cont.]

(b) Use the result of (a)(iii) to answer the following.

(i) Show that colour ratios, such as $R(\mathbf{x})/G(\mathbf{x})$, are invariant to changes in the intensity of the image irradiance at each pixel, but not necessarily to changes in the spectral distribution of $E(\mathbf{x},\lambda)$.

(ii) Define the chromaticity rgb and show that it has similar invariance properties to the colour ratios in (b)(i) above and explain under what condition the chromaticity may also be invariant under changes in the spectral distribution of $E(\mathbf{x},\lambda)$.

(iv) Show that ratios calculated from different pixels $\mathbf{x}$ and $\mathbf{y}$, such as $R(\mathbf{x})/R(\mathbf{y})$, are invariant to changes in the spectral distribution of $E(\mathbf{x},\lambda)$ if such changes are the same at every pixel, but not if the changes are non-uniform (ie vary from pixel to pixel).

(iv) Show that cross-ratios, such as $\dfrac{R(\mathbf{x})}{G(\mathbf{x})}\dfrac{G(\mathbf{y})}{R(\mathbf{y})}$, are invariant both to non-uniform (local) changes in irradiance intensity, and to uniform (global) changes in the spectral distribution of $E(\mathbf{x},\lambda)$.

[10 marks]

(c)

(i) Show that under non-uniform (local) changes in irradiance intensity, and uniform (global) changes in the spectral distribution of $E(\mathbf{x},\lambda)$, each element of the matrix $\mathbf{A}$ in (1) may be written as a product of two factors, one, $S(\mathbf{x})$ say, dependent on pixel location, $\mathbf{x}$, the other dependent on the colour channel.

(ii) Show further that, under the conditions of (c)(i) above, if $\mathbf{C}(\mathbf{x})$ is regarded as an Nx3 matrix over the pixels i=1…N and the RGB colour channels, transformation (1) may be represented as pre-multiplication and post-multiplication of $\mathbf{C}$ by appropriate diagonal matrices.

(iii) Describe Finlayson's comprehensive colour image normalisation algorithm and explain from (c)(ii) above, what kind of colour constancy you would expect it to produce.

[13 marks]

[CONTINUED]

Question 2

(a) Schafer's dichromatic model for the reflectance of light by an object leads to an equation for the colour $\mathbf{C}(\mathbf{x}) = \left(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x})\right)$ at pixel $\mathbf{x}$ in an image of the form

$$\mathbf{C}(\mathbf{x}) = I_S m_b \mathbf{b} + I_S m_s c_s \mathbf{i} \quad . \qquad (2)$$

(i) Explain what each of the terms on the right hand side of equation (2) above stands for and describe how they depend on the scene geometry, spectral content of the illuminant, and on the reflectance properties of the object.

(ii) Describe how, according to Shafer's dichromatic model, you would expect the observed pixel colours to be distributed in the RGB colour cube for images of:

(1) a piece of clothing, such as a man's shirt,

(2) a mirror,

(3) a coloured, glazed ceramic tile.

(iii) Explain what deviations from the distributions in (ii) above you would expect in practice.

[11 marks]

(b) A camera whose colour channels satisfy the integrated white condition is used to derive opponent colour values at each pixel from the difference of the RGB values.

(i) Explain what is meant by the integrated white condition and describe why it is important.

(ii) Show that, according to Shafer's model, if an object is illuminated by white or grey light, the opponent colours obtained from this camera depend on only one of the terms on the right hand side of equation (2) in (a)(i) above.

(iii) Show further that ratios of the above opponent colour values at each pixel should be unchanged under variation of the object's orientation and independent of the brightness of the illuminant.

[9 marks]

[Question 2 cont. over page]

[TURN OVER]

[Question 2 cont.]

(c) The camera described in (b) above is used to take an image of two differently coloured, flat, glazed ceramic tiles illuminated under white light.

(i) Explain how you would use this image to characterise the colour of each tile. What condition must be satisfied if you are to obtain a good characterisation.

(ii) Describe under what circumstances it would be easy to build a machine vision system, utilising the above camera and white light source, to distinguish examples of these two different types of tile on a conveyor belt.

(iii) Explain why the system should, without change, also be able to distinguish unglazed examples of these two types of tile.

(iv) Describe how the system could further be used to distinguish a glazed tile of a particular type from an unglazed tile of the same type. What condition must be satisfied in this case?

[13 marks]

Question 3

(a) A machine vision researcher has developed an image segmentation algorithm for a surveillance system that is based on a model of the image background pixel colour components (R, G, B) being independently distributed as $p_R(R)$, $p_G(G)$ and $p_B(B)$ with means $\overline{R}$, $\overline{G}$, and $\overline{B}$ respectively.

(i) Show that, according to this model, the distribution of background pixel intensities may be obtained from $p_R(R)$, $p_G(G)$ and $p_B(B)$ by two successive convolutions.

(ii) Explain why, if each of the distributions $p_R(R)$, $p_G(G)$ and $p_B(B)$ is assumed to be normal with means $\overline{R}$, $\overline{G}$, $\overline{B}$ and variances σ_R^2, σ_G^2 and σ_B^2, respectively, that the distribution of intensity $I = R + G + B$ will also be normal.

(iii) Evaluate the mean intensity, $\overline{I}$, and the variance, σ^2, of the intensity distribution in (ii) above.

(iv) Comment on the validity, in principle, of using such Gaussian model distributions.

[10 marks]

[Question 3 cont. over page]

[CONTINUED]

[Question 3 cont.]

(b) A second machine vision researcher considers the Gaussian model used in (a) above to be inadequate in practice.

(i) Describe how you would extend the Gaussian model so that it can better describe the distributions of pixel colours likely to be encountered in practice from scenes with a uniform background, such as a uniformly painted wall.

(ii) Describe the method you would use to estimate the parameters of the distribution of pixel colours in this case.

(iii) Illustrate how the method you have described in (b)(ii) above works by using it to estimate the mean, $\bar{I}$, and the variance, σ^2, of the intensity distribution. Comment on the formulae you obtain.

[10 marks]

(c) You are asked to use the system for outdoor surveillance purposes where the background is natural vegetation.

(i) It is often claimed that images of natural scenes are "fractal". Explain what this means and discuss the extent to which such images really are fractal.

(iii) Describe, using the distribution of the pixel intensities as an example, what kind of model you must use to incorporate such effects and indicate how the extent to which such images really are fractal affects the model.

(iii) Describe how you would use the power spectrum of the Fourier transform of the background image intensity to estimate the fractal dimension of your model.

(iv) Explain how you might further extend the distributions you are using in order to enable your model to take into account different aspects of the vegetation, such as leaves, stems and flowers.

[13 marks]

[TURN OVER]

Question 4

(a) A Markov random field $F(m,n)$ used to model image texture is often described by the probability distribution

$$P(\mathbf{F}) = \frac{\exp\left(-\sum_{m,n} U(m,n)\right)}{Z} \quad . \qquad (3)$$

(i) Explain what each of the terms in equation (3) stands for and describe what conditions $P(\mathbf{F})$ and $U(m,n)$ must satisfy in order for $F(m,n)$ to be a Markov random field.

(ii) Describe how the $U(m,n)$ may be represented in terms of interactions between pairs of pixels and over larger pixel cliques.

(iii) Illustrate your answer to (ii) above by means of models designed to encourage smooth variations of the field F, (1) with minimum gradient variation, and (2) with minimum curvature variation.

[10 marks]

(b)

(i) Describe how interactions in the $U(m,n)$ such as those in (1) and (2) in (a)(iii) above can lead to long range correlation in the field F.

(ii) Explain why these interactions make it difficult to calculate a realisation of consistent with the distribution appearing in equation (3) above.

(iii) Describe one method that may be used to calculate a field F consistent with distribution (3).

(iv) Describe what in principle must be done to ensure that the field F obtained by the method you have described in (b)(iii) above does, for example, have minimum gradient variation. Comment on the significance of this requirement in practice.

(v) Explain what you would do in a practical implementation of this technique to overcome the difficulties arising from the requirements in (iv) above.

[13 marks]

[Question 4 cont. over page]

[CONTINUED]

(c)

(i) Explain what is meant by “hidden variables” and describe why you might wish to include such terms in the Markov random field appearing in equation (3) in part (a) above.

(ii) Explain how the hidden variables that you would introduce in (c)(i) affect the interactions between pixels in the Markov random field model you described in (a)(iii)(1) above.

(iii) Describe how the hidden variables you would introduce in c(i) affect each other. Indicate the strength of these interactions relative to each other and to the original interactions in model (a)(iii)(1) above and explain why these interactions should be so.

(iv) Explain why it is difficult to estimate the strengths of these interactions in practice.

[10 marks]

Question 5

(a) A machine vision company wishes to develop a method for efficiently encoding the appearance of an object, such as a human face.

(i) By using the principles of stereo vision, explain why, in principle, it should be possible to synthesize the appearance of the desired object from some particular viewpoint, from a pair of basis images of the object taken from other, different, viewpoints.

(ii) Describe from what range of viewpoints of both the initial and synthesized images such a procedure is likely to be successful, indicating what effects you would expect, in principle, to limit application of the technique.

(iii) Describe why it would be better, if possible, to synthesis the appearance of an object directly from two, initial basis images, rather than by solving the stereo vision problem.

[10 marks]

[Question 5 cont. over page]

[TURN OVER]

[Question 5 cont.]

(b)

(i) Describe briefly, without mathematical details, how the trifocal tensor indicates that it should be possible to synthesize the appearance of an object directly from two initial, basis images.

(ii) Explain what information, in addition to that provided by the trifocal tensor described in (b)(i) above, is required in order to synthesize the new image.

(iii) Use the analogy with stereovision to explain what difficulties you might expect to encounter in using the trifocal tensor to synthesize the new image.

[10 marks]

(c)

(i) Show why, under affine imaging conditions, the trifocal tensor reduces to a set of linear equations.

(ii) Explain how, if the system being developed is to be used for very-low bandwidth face-to-face video communication, you would use a set of control points to estimate the coefficients of the linear equations resulting from (c)(i) above.

(iii) Use the result of (c)(i) above to explain how many coefficients and control points are needed in principle and describe why, in practice, in the application in (c)(ii) above you would use: (1) a few more coefficients, and (2) many more control points.

(iv) Describe how you would utilise the coefficients used in (c)(iii) above to help complete synthesis of the desired image at the receiver in this video communication system. Indicate briefly any remaining problems that would have to be overcome and how you might do so.

[13 marks]

[END OF PAPER]