

UNIVERSITY OF TORONTO  
Department of Computer Science

DECEMBER 2004 EXAMINATIONS

CSC487/2503: Foundations of Computer Vision

Duration: 3 hours

Aids allowed: None

There are **15** pages total (including this page)

Please answer questions on the exam pages in the space provided. Only use the backs of pages if necessary. There are **6** main questions, most of which have multiple parts. Read questions carefully and answer as neatly, clearly, and concisely as possible; you will be marked on the clarity and correctness of your answers. Should a question be unclear or ambiguous, make a *reasonable* interpretation and state what you have assumed before answering.

Family name: \_\_\_\_\_

Given names: \_\_\_\_\_

Student number: \_\_\_\_\_

Question	Marks
1	_____/30
2	_____/12
3	_____/14
4	_____/14
5	_____/17
6	_____/14
Total	_____/101

**1. Short answer questions [30 marks]**

**(a) - 4 marks** Specify the mathematical form of a 2D homography. What classes of surfaces and camera motions produce image homographies between two or more views?

**(b) - 5 marks** Define radiance and irradiance and specify the units in which they are expressed. Define the bidirectional reflectance distribution function and specify how it is used to specify the light reflected from a surface toward the camera.

**(c) - 4 marks** Write the integral for the convolution of two 1D functions  $f_1(x)$  and  $f_2(x)$ ? Express this operation in terms of the Fourier transforms of the two functions,  $\hat{f}_1(\omega)$  and  $\hat{f}_2(\omega)$ ?

**(d) - 5 marks** In words (and figures if useful), explain the underlying signal model and assumptions, as well as the three objective criteria, that are the basis for the 1D Canny edge detector?

**(e) - 4 marks** What is meant by *coarse-to-fine* optical flow estimation, and under what conditions is a coarse-to-fine scheme useful?

**(g) - 4 marks** What does *IT* stand for in IT-search? And briefly explain what IT-search is.

**(f) - 4 marks** What is the objective function for Ncut image segmentation?

**2. Multiscale image transforms [12 marks]**

Assuming that images  $I(x)$  are represented as points  $\mathbf{I}$  in a vector space, we can express up-sampling, down-sampling, and smoothing as matrix operations.

**(a) - 4 marks** Given a simple depiction of the form of matrices for up-sampling, down-sampling, and smoothing.

**(b) - 5 marks** In terms of these matrices, show the recursive equations for the construction of the Laplacian pyramid.

**(c) - 3 marks** From the Laplacian pyramid coefficients, write the recursive equations that can be used to reconstruct the original image.

### 3. Robust parameter estimation [14 marks]

Let's say we wish to fit a 1D curve  $c(t) = a_0 + a_1t + a_2t^2$  to a set of  $N$  observed 1D image locations,  $\{x_j\}_{j=1}^N$ , in frames  $t_j$ , for  $j = 1 \dots N$ . That is, you are given  $\{(x_j, t_j)\}_{j=1}^N$ . Further assume that we measure error between the curve and the measurement as  $\rho(c(t_j) - x_j)$  for a robust estimator  $\rho$ :

**(a) - 3 marks** Define the *influence function* and the *weight function* in terms of a general robust estimator  $\rho(e)$ .

**(b) - 2 marks** Using this function,  $\rho(e)$ , write an objective function that you might try to minimize in order to find  $c(t)$ .

**(c) - 3 marks** Specify the mathematical form of a robust estimator function  $\rho(e)$  that you might use if you are told that the data set has outliers.

**(d) - 6 marks** For the specific robust estimator in Part (c) above, derive the weights for and specify the form of an iteratively reweighted least squares solution.

**4. Optical flow estimation [14 marks]**

**(a) - 2 marks** State and briefly explain the motivation for the *brightness constancy assumption*?

**(b) - 3 marks** Given the brightness constancy assumption, derive the form of a *gradient constraint equation* for image motion estimation.

**(c) - 6 marks** Assume that the image is the sum of two drifting sinusoidal signals with spatial and temporal frequencies  $\mathbf{k}_j$  and  $\omega_j$  for  $j = 1, 2$ , and amplitudes  $a_j$ :

$$I(\mathbf{x}, t) = a_1 \sin(\mathbf{k}_1 \cdot \mathbf{x} + \omega_1 t) + a_2 \sin(\mathbf{k}_2 \cdot \mathbf{x} + \omega_2 t)$$

where  $\mathbf{x}$  and  $t$  denote 2D spatial position and time, and  $\mathbf{k} \cdot \mathbf{x}$  denotes the usual dot product.

Formulate the brightness constancy equations in order to estimate constant image translation  $\mathbf{v}$  within a region containing pixels  $\{\mathbf{x}_n\}_{n=1}^N$ . Place the equations in matrix form, giving names to the matrices, and give 1 line of Matlab code to solve for the velocity.

**(d) - 3 marks** Specify under what conditions on the sinusoidal parameters,  $a_1, a_2, \mathbf{k}_1, \mathbf{k}_2, \omega_1, \omega_2$ , the aperture problem occurs and one cannot uniquely constrain the optical flow.

## 5. Epipolar geometry and camera calibration [17 marks]

Assume you are given  $K$  3D points  $\{\mathbf{X}_k\}_{k=1}^K$ , that project to 2D pixel coordinates  $\{\mathbf{x}_k^L\}_{k=1}^K$  and  $\{\mathbf{x}_k^R\}_{k=1}^K$  in left and right images of a stereo image pair. The fundamental matrix  $F$  captures both intrinsic and extrinsic information about the stereo camera geometry.

**(a) - 1 marks** How many corresponding points are required for Hartley's eight-point algorithm?

**(b) - 6 marks** Derive the fundamental matrix in terms of the intrinsic and extrinsic calibration matrices for the left and right cameras.

**(c) - 2 marks** What is the maximum rank of the fundamental matrix?

**(d) - 3 marks** How many essential degrees of freedom (unknowns) are there in the fundamental matrix? Briefly explain your answer.

**(e) - 2 marks** Briefly define the term *epipolar line*, and explain the relation between epipolar lines and the fundamental matrix.

**(f) - 3 marks** The change in coordinates from one camera to another can be expressed as a translation of the camera center and a rotation of the image plane. Show that epipolar lines obtained from the fundamental matrix do not depend on the magnitude or sign of the translation between camera centers.

## 6. Orthographic factorization [14 marks]

Suppose you are given a collection of  $K$  2D image points in  $F$  frames, i.e.,  $\{\mathbf{x}_{k,f}\}$  for  $k = 1 \dots K$  and  $f = 1 \dots F$ . Assume that the data are organized so that for each  $k$ , the points  $\{\mathbf{x}_{k,f}\}$  for  $f = 1 \dots F$  are scaled orthographic projections of a 3D point  $\mathbf{X}_k$ .

**(a) - 3 marks** Express the scaled orthographic projection of  $\mathbf{X}_k$  onto  $\mathbf{x}_{k,f}$  in terms of a projection matrix  $\mathbf{M}_f$ . Express  $\mathbf{M}_f$  in terms of intrinsic and extrinsic camera parameters.

**(b) - 4 marks** Write the form of the data matrix and its relation to a factorization in terms of shape and image projection matrices. What is the rank of the data matrix? Explain your answer.

**(c) - 3 marks** Explain what is meant by affine reconstruction of shape, and explain the mathematical basis for the affine ambiguity.

**(d) - 2 marks** Explain how properties of the projection matrix  $M_f$  can be used to resolve the affine ambiguity.

**(e) - 2 marks** What are the remaining ambiguities in the reconstruction?

**END OF EXAM**