Engineering Tripos Part IIB FOURTH YEAR

Module 4F12: Computer Vision

Examples Paper 1

Straightforward questions are marked †
Tripos standard (but not necessarily Tripos length) questions are marked $*$

1. † *Images*

   Images are stored as pixel arrays of quantised intensity values. Typically each pixel has a brightness value in the range 0 (black) to 255 (white), and is stored as a single byte (8 bits). Compute the storage requirements (in bytes per second) for a stereo pair of HD video cameras grabbing grey-level images of size 1920×1080 pixels at 25 frames per second. Approximately how many pages of text require the same amount of storage as one second of stereo video?

2. $*$ *Smoothing by convolution with a Gaussian*

   A commonly used 1D smoothing filter is the Gaussian:

$$g_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

   where σ determines the size of the filter. Show that repeated convolutions with a series of 1D Gaussians, each with a particular standard deviation σ_i, is equivalent to a single convolution with a Gaussian of variance $\sum_i \sigma_i^2$.

3. *Generating the Gaussian filter kernel*

   A discrete approximation to a 1D Gaussian can be obtained by sampling the function $g_\sigma(x)$. In practice, samples are taken uniformly until the truncated values at the tails of the distribution are less than 1/1000 of the peak value.

   (a) For $\sigma = 1$, show that the filter obtained in this way has a size of 7 pixels and coefficients given by:

0.004	0.054	0.242	0.399	0.242	0.054	0.004

   What property of the coefficients ensures that regions of uniform intensity are unaffected by smoothing?

   (b) Using the same truncation criterion, what would be the size of the discrete filter kernel for $\sigma = 5$? Show that, in general, the size of the kernel can be approximated as $2n + 1$ pixels, where n is the nearest integer to $3.7\sigma - 0.5$.

   (c) The filter is used to smooth an image as part of an edge detection procedure. What factors affect the choice of an appropriate value for σ?

4. † *Discrete convolution*

The following row of pixels is smoothed with the discrete 1D Gaussian kernel given in question 3(a) ($\sigma = 1$). Calculate the smoothed value of the pixel with intensity 118.

46	45	45	48	50	53	55	57	77	99	118	130	133	134	133	132	132	132	133

5. *Derivative of convolution theorem*

(a) Show that smoothing an intensity signal with a Gaussian and then differentiating the smoothed signal is equivalent to convolution with the derivative of a Gaussian:

$$\frac{d}{dx}[g_\sigma(x) * I(x)] = g'_\sigma(x) * I(x)$$

where $g'_\sigma(x)$ is the first derivative of the Gaussian function.

(b) Hence, or otherwise, show how "edges" in an intensity function $I(x)$ can be localised at the zero-crossings of $g''_\sigma(x) * I(x)$, where $g''_\sigma(x)$ is the second derivative of the Gaussian function.

6. *Differentiation and 1D edge detection*

Show how an approximation to the first-order spatial derivative of $I(x)$ can be obtained by convolving samples of $I(x)$ with the kernel

1/2	0	-1/2

The smoothed row of pixels in question 4 is shown below.

x	x	x	48	50	53	56	64	79	98	115	126	132	133	133	132	x	x	x

Find the first order derivatives and localise the intensity discontinuity.

7. *Decomposition of 2D convolution*

Smoothing a 2D image involves a 2D convolution with a 2D Gaussian:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp - \left(\frac{x^2 + y^2}{2\sigma^2}\right)$$

Show that this can be performed by two 1D convolutions: i.e.

$$G_\sigma(x, y) * I(x, y) = g_\sigma(x) * [g_\sigma(y) * I(x, y)]$$

What is the advantage of performing two 1D convolutions instead of a 2D convolution?

8. $*$ *Isotropic and directional edge finders*

   The Marr–Hildreth operator convolves the image with a discrete version of the Laplacian of a Gaussian and then localises edges at the resulting zero-crossings. Show that the Laplacian of a Gaussian is an isotropic (ie. rotationally symmetric) operator. Hence explain why the operator produces zero-crossings along an ideal step edge.

   The Canny operator is a directional edge finder. It first localises the orientation of the edge by computing

$$\hat{\mathbf{n}} = \frac{\nabla\left(G_\sigma(x,y) * I(x,y)\right)}{\left|\nabla\left(G_\sigma(x,y) * I(x,y)\right)\right|}$$

   and then searches for a local maximum of $|\nabla(G_\sigma * I)|$ in the direction $\hat{\mathbf{n}}$. Show that this is equivalent to finding zero-crossings in the directional second derivative of $(G_\sigma * I)$ in the direction $\hat{\mathbf{n}}$, ie. finding zero crossings in

$$\frac{\partial^2 (G_\sigma * I)}{\partial s^2}$$

   where s is a length parameter in the direction $\hat{\mathbf{n}}$.

   What are the advantages and disadvantages of isotropic and directional operators?

9. $*$ *Auto-correlation and corner detection*[Tripos 2012]

   (a) Show that the weighted sum of squared differences (SSD) between a patch (window W) of pixels in image $I(x,y) = I(\mathbf{x})$ and another patch of pixels taken by shifting the window by a small amount in the direction $\mathbf{n}$ can be expressed approximately by:

$$C(\mathbf{n}) = \sum_{\mathbf{x}\in W} w(\mathbf{x})(I(\mathbf{x}+\mathbf{n}) - I(\mathbf{x}))^2 \approx \sum_{\mathbf{x}\in W} w(\mathbf{x}) I_n^2$$

   where $I_n = \nabla I(\mathbf{x}).\mathbf{n}$.

   (b) Hence show that the weighted SSD can be represented by:

$$C = \mathbf{n}^T \mathrm{A}\mathbf{n}$$

   where A is matrix of smoothed intensity gradients (sometimes called a second-moment or autocorrelation matrix) defined as follows:

$$\mathrm{A} \equiv \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \left\langle I_y^2 \right\rangle \end{bmatrix}$$

   where $I_x \equiv \partial I/\partial x$, $I_y \equiv \partial I/\partial y$ and $\langle\rangle$ denotes a 2-dimensional weighting (smoothing) operation.

   (c) How are the directional derivatives computed? How are the 2D weighted (smoothed) values obtained?

(d) Show how A can be analysed to detect corner features and give details of the Harris-Stephens corner detection algorithm.

(Note — Part IA Maths revision) For a real, symmetric $n \times n$ matrix A the minimum and maximum values of

$$C = \frac{\mathbf{n}^T \mathrm{A}\mathbf{n}}{\mathbf{n}^T \mathbf{n}}$$

are given by

$$\lambda_1 \leq C \leq \lambda_n$$

where λ_1 and λ_n are the minimum and maximum eigenvalues of A respectively.

10. $*$ *Feature detection and scale space* [Tripos 2011]

Consider an algorithm to detect interest points (features of interest) in a 2-D image for use in matching.

(a) Show how different resolutions of the image can be represented efficiently in an *image pyramid*. Your answer should include details of the implementation of smoothing within an octave and subsampling of the image between octaves.

(b) How can *band-pass* filtering at different scales be implemented efficiently using the image pyramid? Show how image features such as *blob-like* shapes can be localized in both position and scale using band-pass filtering.

(c) Explain how interest points in different images can be matched. Give details of 3 suitable descriptors.

Suitable past Tripos questions: Q1 on all exams 1996-2015

Roberto Cipolla
October 2015