Paper 8 Information Engineering Elective: Image Search

Examples Paper

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 cameras grabbing grey-level images of size 512×512 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. *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 | -1 |

   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.

6. *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?

7. *Correlation and Convolution*

   The correlation of a template $g(x, y)$ (perhaps taken from a different image) and the image $I(x, y)$, we may use the *cross-correlation* which is defined by:

$$c(x, y) = \int \int I(u, v) g(u + x, v + y) du dv$$

   This is usually normalised (by the root-mean-square intensities of each region) so that the cross-correlation will take its maximum value of 1.0 when the 2 signals are identical.

   Both cross-correlation and *convolution* involve shifting, multiplication and summation operations. Write down the equation describing the convolution of two signals, $I(x, y)$ and $g(x, y)$. What are the differences between the cross-correlation and convolution of two signals?

8. $*$ *Corner detection*

(a) (Revision — Part IA Maths) For a real, symmetric $n \times n$ matrix A show that 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.

(b) A matrix of smoothed intensity gradients is 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 smoothing operation. Show how A can be analysed to detect corner features. How are the directional derivatives computed? How are the smoothed values obtained?

9. $*$ *Texture* What is meant by image texture? Show how image texture can be characterized by the output of a bank of filters. Give details of typical filter kernals used and the convolution operations required.

10. $*$ *Feature descriptors* Describe 3 different descriptors for interest points and comment on their relative advantages and disadvantages when used for image matching.

Answers

2. 1.3×10^7 Bytes/s; ≈ 3000 pages
4. (b) 37 pixels.
5. 115 (to the nearest integer)
7. Between the pixel with intensity 79 and the pixel with intensity 98. More precisely, two-thirds of the way.

Roberto Cipolla
October 2005