Computer Vision · Assignment 1: Foundations

AIMS Cameroon · Structured Master's, Data Science · 2026–2027

Individual work · 100 points · Due: start of Week 2, Day 3 (set the exact date on the course page)

Submission. One PDF typeset in LaTeX (handwritten scans are not accepted) plus one Jupyter notebook that reproduces every figure and number in Problem 5. Name the files `A1_<surname>.pdf` and `A1_<surname>.ipynb`. You may discuss ideas with classmates, but the write-up and code must be your own. Cite any source you use, including AI assistants, and state what you used them for.

Problem 1: Quantisation noise (12 points)

An image with values in $[0, 256)$ is quantised uniformly to $L = 2^b$ levels with step $\Delta = 256/L$. Model the quantisation error e as uniform on $[-\Delta/2, \Delta/2]$.

(a) (4) Show that $\mathbb{E}[e] = 0$ and $\mathrm{Var}(e) = \Delta^2/12$.

(b) (5) Using $\mathrm{PSNR} = 10 \log_{10}(255^2/\mathrm{MSE})$, show that $\mathrm{PSNR} \approx 6.02\, b + 10.8$ dB.

(c) (3) Using the PSNR-versus-bits curve from the Day 2 tutorial notebook, comment on how well this model fits for $b = 1$ and for $b = 8$, and explain any discrepancy.

Problem 2: Properties of convolution (20 points)

Let $(f * h)[n] = \sum_k f[k]\, h[n-k]$ in 1D and analogously in 2D.

(a) (4) Prove that discrete convolution is commutative: $f * h = h * f$.

(b) (4) Prove that it is shift-equivariant: if $g = f * h$ and $f_\tau[n] = f[n-\tau]$, then $(f_\tau * h)[n] = g[n-\tau]$.

(c) (4) Show that the 2D Gaussian kernel is separable, $G_\sigma(x, y) = g_\sigma(x)\, g_\sigma(y)$, and compute the number of multiplications per pixel for a $K \times K$ kernel applied directly versus separably. Evaluate for $K = 31$.

(d) (5) Show (in the continuous setting, by any method) that $G_{\sigma_1} * G_{\sigma_2} = G_{\sqrt{\sigma_1^2 + \sigma_2^2}}$. Explain what this implies for building a Gaussian scale space incrementally (as in SIFT).

(e) (3) Two stacked 3×3 convolutions are equivalent to one convolution with a kernel of what size? Use associativity to justify your answer.

Problem 3: Histogram equalisation (13 points)

(a) (5) Let r be a continuous random variable with strictly increasing CDF F. Prove that $U = F(r)$ is uniform on $[0, 1]$. Where does the proof use "strictly increasing"?

(b) (4) A 4×4 image with $L = 8$ has pixel values $\begin{smallmatrix} 2 & 2 & 3 & 3 \\ 2 & 3 & 3 & 4 \\ 3 & 3 & 4 & 4 \\ 3 & 4 & 4 & 5 \end{smallmatrix}$. Compute the histogram, the equalisation map s_k and the equalised image.

(c) (4) Explain why equalising R, G and B independently distorts colours, and describe the standard alternative.

Problem 4: The Harris detector (15 points)

(a) (5) Starting from $E(u, v) = \sum_W w(x, y)[I(x+u, y+v) - I(x, y)]^2$, derive $E(u, v) \approx [u\ v]\, M\, [u\ v]^\top$ and state M.

(b) (4) Show that $R = \det M - k(\operatorname{tr} M)^2 = \lambda_1\lambda_2 - k(\lambda_1 + \lambda_2)^2$, and sketch the regions of the (λ_1, λ_2) plane where $R > 0$, $R < 0$ and $|R|$ is small.

(c) (3) If the image is rotated by an angle α, the gradients rotate: $\nabla I' = Q\nabla I$ with Q a rotation matrix. Show that $M' = QMQ^\top$ and conclude that the Harris response is rotation invariant.

(d) (3) Is R invariant to an intensity change $I \mapsto aI + c$? To a change of image scale? Justify both answers.

Problem 5: Mini-project, a classical vision pipeline on your own data (40 points)

Collect **your own images** with a phone camera (at least 30 photos). Choose *one* of the two tracks below and write a report section of about 3 pages with figures.

Track A: Canny from scratch. Implement, using only NumPy (and SciPy for `ndimage.label`): separable Gaussian smoothing, Sobel gradients, non-maximum suppression with direction quantisation, and double threshold + hysteresis.

- Compare your edges with `cv2.Canny` on at least 5 of your photos using a tolerant agreement metric (precision/recall within ± 1 pixel, as in the Day 4 tutorial notebook). Report a table.
- Study the effect of σ and of the thresholds on two photos. Show a 3×3 grid of results and discuss it.
- Evaluate robustness: add Gaussian noise at three levels and plot agreement against noise level, with and without a median pre-filter.

Track B: Engineered features for recognition. Photograph at least 3 classes of objects relevant to your context (e.g. three crop leaves, three kinds of coins or banknotes, three fruits), at least 10 images per class under varied lighting and backgrounds.

- Build and compare at least three representations: raw downsampled pixels, a colour histogram (in a colour space you justify), and HOG. Use the same classifier (e.g. logistic regression or a linear SVM) with cross-validation.
- Report per-class precision and recall and a confusion matrix, not only accuracy.
- Test robustness to one nuisance factor of your choice (brightness, rotation, shift, scale) applied to the test set, and explain the results using the design of each feature.

Marking (both tracks).

Correct, readable implementation (code runs top to bottom)	12
Experimental design (controls, fair comparison, honest metrics)	10
Figures and tables (labelled, readable, referenced in the text)	8
Discussion linking results to the theory from lectures	8
LaTeX presentation and references	2
Total	40