Lecture 9:

Diffusion Filtering: Parameter Selection

Contents

Motivation

- Diffusion filters are flexible tools that may give good results.
- However, they require to specify several parameters.
- Non-experts often regard this as a difficult problem and ask for
  - guidelines for parameter selection
  - algorithms that select the parameters automatically.
- Several reasonable strategies exist, but no universal guidelines.
  No algorithm can tell us what we really want to find in an image:

"One man's signal is another man's noise."

Selection of the Stopping Time

Stopping Time and Scale

- The stopping time is a natural scale parameter in all diffusion methods.
- For *linear* diffusion filtering, it can be related to the standard deviation σ of Gaussian convolution (Lecture 2):

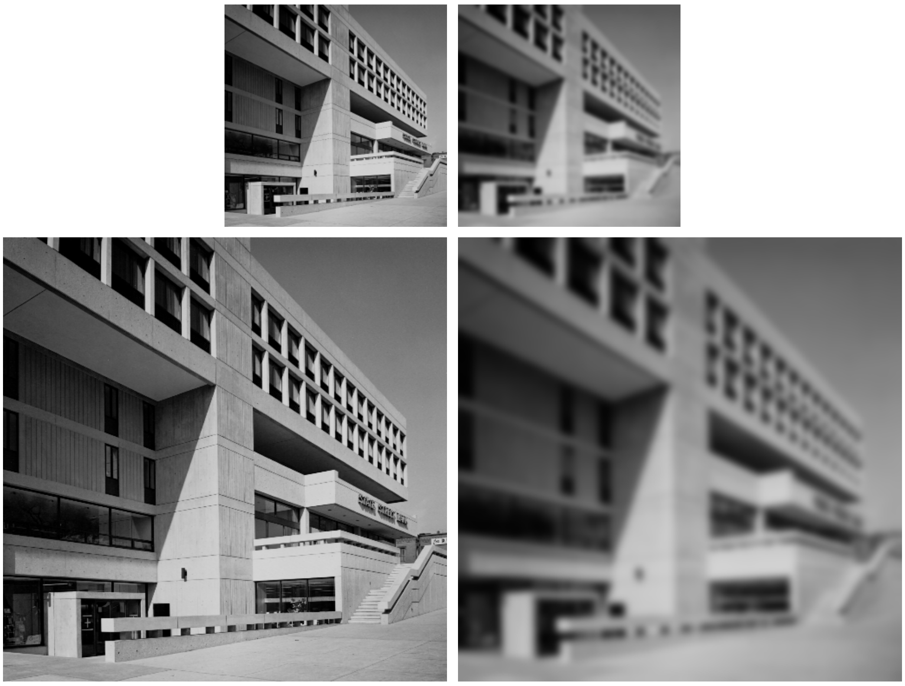
$$T = \frac{1}{2}\sigma^2.$$

  This gives an intuitive relation between diffusion time and spatial extension of the smoothing scale.
- For *nonlinear* diffusion filters, such a uniform scale interpretation does not exist: Small-scale high contrast details can have a longer lifetime than large-scale low contrast structures.

How Does the Stopping Time Scale Under Resampling?

- In theory, the stopping time is a parameter of the continuous diffusion model that has nothing to do with the discretisation.
- In practice, many image processing people assign unit length to a pixel.
  In this case, doubling the pixel number in each direction means enlarging the image domain Ω by a factor 2 in each direction.
- For linear diffusion filters, $T = \sigma^2/2$ implies that doubling the spatial smoothing scale σ requires to enlarge the diffusion time by a factor 4.
- Thus, in the discrete setting, we need 4 times as many iterations.
- Moreover, one explicit iteration step requires a complexity that is linear in the pixel number (may be higher for (semi-)implicit methods if the linear system cannot be solved in linear complexity).
- In 2-D, doubling the pixel number per direction increases the computing time by a factor 4 (factor 8 in 3-D).
- Thus, the total computational effort to obtain a comparable smoothing effect increases by a factor 16 (factor 32 in 3-D) !!

Selection of the Stopping Time (3)



(a) **Top left:** Original image with lower resolution, $\Omega = (0,256) \times (0,256)$. (b) **Top right:** Linear diffusion filtering of (a) with stopping time $T = 2.5$. (c) **Bottom left:** Original image with higher resolution, $\Omega = (0,512) \times (0,512)$. (d) **Bottom right:** Linear diffusion filtering of (c) requires now $T = 10$ to give comparable results. **Author:** J. Weickert (2007).

MIA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Selection of the Stopping Time (4)

Stopping Time Selection for Nonlinear Diffusion: Ideas That Fail

- **Specifying a Smoothing Scale**

  We have seen that specifying a spatial smoothing scale σ does not work:
  No nice scaling behaviour as for linear diffusion ($T = \sigma^2/2$).

- **Using Linear Diffusion Estimates**

  If the diffusivity is ≤ 1, we may use linear diffusion to get a lower estimate:

$$T \geq \frac{1}{2}\sigma^2.$$

  However, this estimate can be far too small:
  Nonlinear diffusion can be much slower than linear, in particular if the contrast parameter λ is too small.

- **Using Ad Hoc Simplification Criteria**

  Example: number of extrema should be reduced by a factor 3.
  This may not work either:
  For dimensions ≥ 2, diffusion filters may even create extrema (cf. Lecture 3).
  Are there better simplification criteria?

MIA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Selection of the Stopping Time (5)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Effect of using identical stopping times for linear and nonlinear diffusion. (a) **Top left:** Original image, $\Omega = (0, 256) \times (0, 256)$. (b) **Top right:** Linear diffusion filtering of (a) with stopping time $T = 20$. (c) **Bottom left:** Perona-Malik diffusion, $\lambda = 3$, $T = 20$. (d) **Bottom right:** Perona-Malik diffusion, $\lambda = 1$, $T = 20$. **Author:** J. Weickert (2008).

Selection of the Stopping Time (6)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Lyapunov Criterion (Weickert 1999)

- We do already know a large family of simplification measures that are not ad hoc, namely the Lyapunov functionals (Lectures 4 and 8)

$$\Phi(u(.,t)) := \int_\Omega r(u(\mathbf{x},t))\, d\mathbf{x}$$

  for some convex C^2 function r.

- Since $\Phi(u(.,t))$ is decreasing in t and bounded from below by $\Phi(\mu)$, where μ is the average grey value of f, we know that

$$\Psi(u(.,t)) := \frac{\Phi(f) - \Phi(u(.,t))}{\Phi(f) - \Phi(\mu)}$$

  increases in a monotone way from 0 to 1.

- Specifying a certain simplification by prescribing a value for Ψ gives us a simple and robust *a posteriori* stopping criterion.

Selection of the Stopping Time (7)

The Lyapunov criterion replaces one parameter selection problem by another one. So what has been gained?

- robustness, since it avoids a number of problems:
  - recalculation of the stopping time when the image is resampled
  - strong influence of the contrast parameter on the stopping time
- has been used successfully when giving diffusion programmes to non-experts in image processing, e.g. clinicians

There are many Lyapunov functionals. Which one should I take?

- Use a simple one such as the variance

$$\mathrm{var}(u) \;:=\; \frac{1}{|\Omega|}\int_\Omega (u(\mathbf{x},t)-\mu)^2\,d\mathbf{x}$$

where $\mu := \frac{1}{|\Omega|}\int_\Omega f(\mathbf{x})\,d\mathbf{x}$ is the average grey value.

Selection of the Stopping Time (8)

Effect of using identical variances for linear and nonlinear diffusion. (a) **Top left:** Original image, $\Omega = (0,256)\times(0,256)$, variance: 2449. (b) **Top right:** Linear diffusion filtering of (a) with stopping time $T = 0.72$ yielding a variance of 2306. (c) **Bottom left:** Perona-Malik diffusion with $\lambda = 3$, $T = 5$, variance 2306. (d) **Bottom right:** Perona-Malik diffusion with $\lambda = 1$, $T = 143.75$, variance 2306. **Author:** J. Weickert (2008).

M I A
1 2
3 4
5 6
7 8
9 10
11 12
13 14
15 16
17 18
19 20
21 22
23

Decorrelation Criterion (Mrázek / Navara 2003)

- *a posteriori* strategy without introducing new parameters
- based on two assumptions:
  - signal and noise are uncorrelated
  - iterative filter that removes noise first
- algorithm:
  - Compute the correlation coefficient between the "signal" $u(\mathbf{x}, t)$ and the "removed noise" $f(\mathbf{x}) - u(\mathbf{x}, t)$ for all times (iteration steps).
  - For the optimal stopping time, this correlation becomes minimal.
- works well in practice for a number of iterative denoising methods
- no theoretical guarantee that there is a unique minimum

M I A
1 2
3 4
5 6
7 8
9 10
11 12
13 14
15 16
17 18
19 20
21 22
23

How is the Correlation Computed?

- The *mean* or *(empirical) expectation* of a discrete signal $\mathbf{f} \in \mathbb{R}^N$ is given by

$$\bar{\mathbf{f}} := E(\mathbf{f}) := \frac{1}{N} \sum_{i=1}^{N} f_i.$$

- The *variance* of $\mathbf{f}$ is defined as

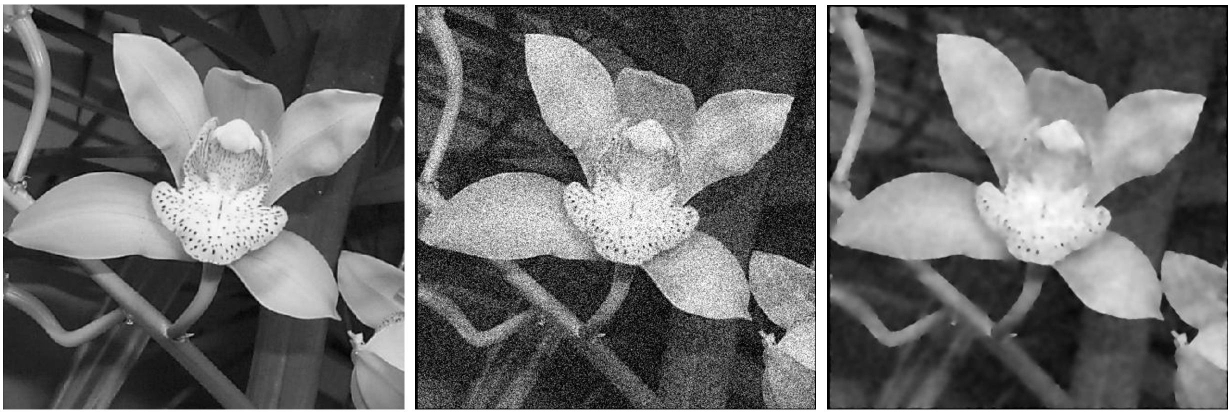
$$\mathrm{var}(\mathbf{f}) := E((\mathbf{f} - \bar{\mathbf{f}})^2),$$

and the *covariance* of two signals $\mathbf{f}$, $\mathbf{g}$ is given by

$$\mathrm{cov}(\mathbf{f}, \mathbf{g}) := E((\mathbf{f} - \bar{\mathbf{f}})(\mathbf{g} - \bar{\mathbf{g}})).$$

- The *correlation coefficient* of $\mathbf{f}$ and $\mathbf{g}$ is computed as

$$\mathrm{corr}(\mathbf{f}, \mathbf{g}) := \frac{\mathrm{cov}(\mathbf{f}, \mathbf{g})}{\sqrt{\mathrm{var}(\mathbf{f})}\,\sqrt{\mathrm{var}(\mathbf{g})}}.$$



Application of the decorrelation criterion. (a) **Left:** Original image. (b) **Middle:** Noisy version. (c) **Right:** Denoised with a nonlinear diffusion filter stopped according to the decorrelation criterion. **Authors:** P. Mrázek and M. Navara (2003).

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Selection of the Contrast Parameter

The Contrast Parameter Selection Problem

- Most nonlinear diffusion filters require to specify a contrast parameter λ.
- Example: Perona-Malik filter with diffusivity

$$g(|\nabla u|^2) \;=\; \frac{1}{1+|\nabla u|^2/\lambda^2}$$

- Contrast parameter distinguishes between high-contrast areas (edges) that should be preserved and low-contrast areas (interior of regions) that should be smoothed.
- Problem: Optimal value for λ depends strongly on the image, in particular its grey value range.

M I A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Quantile Criterion (Perona / Malik 1990)

- Compute a cumulative histogram of the initial image gradient $|\nabla f|$ (for Perona-Malik) or $|\nabla f_\sigma|$ (for regularised Perona-Malik or EED)
- Set λ to a specified quantile of the histogram.
- Example:
  - The 70 % quantile is the parameter that separates the histogram into 70 % smaller values and 30 % larger values.
  - In this case we assume that 30 % of all pixels belong to edge regions.
- There are efficient algorithms available for computing the quantile; see e.g. Press et al. (Numerical Recipes)
- Same effect as the with the Lyapunov criterion:
  - replaces one parameter selection problem by another
  - more robust than specifying a certain value for λ, since a default quantile value is invariant under linear grey value rescalings
  - good experiences with non-experts since the quantile is very intuitive

Selection of Other Parameters

Noise Scale σ

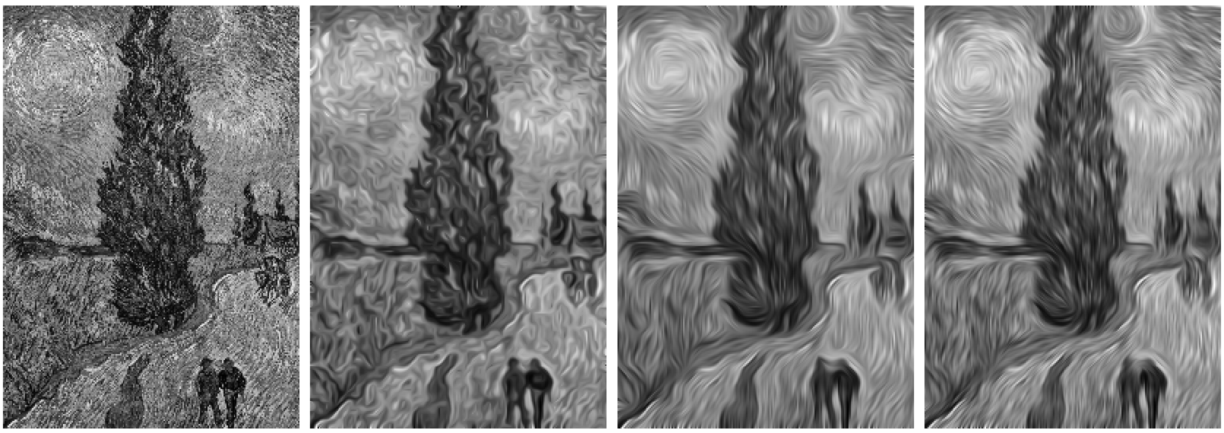
- required when using regularised Perona-Malik diffusion, EED or CED
- determines the size of details that are regarded as unimportant or noise
- If pixels have unit size, than typical values for σ are in the range from 0.5 to 3, depending on the noise level
- Use the smallest σ that works:
  - If σ is too small, noise and unimportant structures may survive too long.
  - If σ is too large, important image structures are destroyed.

Selection of Other Parameters (2)

Integration Scale ρ

- required for CED
- should be larger than the distance of two adjacent lines
- typically significantly larger than the noise scale: $\rho \geq 2\sigma$
- Overestimations are less critical than underestimations.
- Often a default value of $\rho \approx 5$ works well.

Selection of Other Parameters (3)



Impact of the integration scale on coherence-enhancing anisotropic diffusion ($\sigma = 0.5$, $t = 8$). (a) **Left:** "Lane under Cypresses below the Starry Sky" by van Gogh (Auvers-sur-Oise, 1890; Otterlo, Rijksmuseum Kröller–Müller), $\Omega = (0, 203) \times (0, 290)$. (b) **Middle left:** Filtered with $\rho = 1$. (c) **Middle right:** $\rho = 4$. (d) **Right:** $\rho = 6$.

Methods with Fewer Parameters

- A very interesting nonlinear diffusion filter is the *total variation diffusion (TV flow)*

$$\partial_t u \;=\; \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$$

  (Andreu et al. 2000, 2002)

- as simple as linear diffusion:
  no contrast parameter λ, no noise scale σ, no integration scale ρ
- represents borderline between forward and backward diffusion:
  1-D flux $\phi(s) = \mathrm{sgn}(s)$ is constant for $s > 0$
- In 1D, the contrast reduces linearly in time.
- finite (!) extinction time
- numerically more difficult because of the unbounded diffusivity $g(|\nabla u|^2) = \frac{1}{|\nabla u|}$
  (more details in a later lecture)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Evolution under TV diffusion. (a) **Top left:** Original image with Gaussian noise. (b) **Top right:** After applying TV diffusion with $T = 50$. (c) **Bottom left:** $T = 200$. (d) **Bottom right:** $T = 3000$. This is already beyond the extinction time where the image becomes constant. **Author:** J. Weickert (2006).

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Summary

- The stopping time selection is the main parameter selection problem for nonlinear diffusion filtering.
- For linear diffusion filtering, the diffusion time should be related to the square of the desired spatial smoothing scale.
- In 2-D, using twice as many pixels in each direction increases the computational burden by a factor 16.
- Lyapunov functionals such as the variance replace the stopping time for nonlinear diffusion filters by a more robust measure.
- The decorrelation criterion is simple and does not require any specific knowledge about the noise (apart from being uncorrelated with the signal).
- The contrast parameter may be selected according to the quantile criterion.
- Noise scale or integration scale have an intuitive interpretation.
  In practice, their specification is less critical.
- It may be useful to consider PDEs with fewer parameters such as TV diffusion.

References

- J. Weickert: Coherence-enhancing diffusion of colour images. *Image and Vision Computing*, Vol. 17, 201-212, 1999.
  (Lyapunov criterion)
- J. Weickert: Design of nonlinear diffusion filters. In B. Jähne, H. Haußecker (Eds.): *Computer Vision and Applications*. Academic Press, San Diego, pp. 439-458, 2000.
  (general discussion on parameter selection)
- P. Mrázek, M. Navara: Selection of optimal stopping time for nonlinear diffusion filtering. *International Journal of Computer Vision*, Vol. 52, No. 2/3, pp. 189–203, 2003.
  `http://www.mia.uni-saarland.de/publications.shtml`
  (correlation criterion for stopping time selection)
- P. Perona, J. Malik: Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 629–639, 1990.
  (quantile criterion for selecting the contrast parameter)
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery: *Numerical Recipes in C*. Second edition, Cambridge University Press, 1992. `http://www.nrbook.com/a/bookcpdf.php`
  (fast algorithms for computing a quantile in a histogram)
- F. Andreu–Vaillo, V. Caselles, J. M. Mazón: *Parabolic Quasilinaer Equations Minimizing Linear Growth Functionals*. Birkhäuser, Basel, 2004.
  (analysis of total variation diffusion)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

Announcement

- Those who study mathematics or computer science, please make sure that you register your participation in this class via the HISPOS system by May 14.
- All students who want to participate in the exam must also have registered via our web site.