Structure Preserving Image Smoothing

Erkut Erdem
Today

- What is an image? What is image smoothing?
- A Little Bit of History
- Why is image smoothing interesting? A Personal Answer
- What I’ve done
  - Context-guided image filtering (joint work with Sibel Tari)
  - Structure Preserving Image Smoothing via Region Covariances (joint work with Levent Karacan, Aykut Erdem)
- Where we’re going
A little bit of history

- Gaussian Filtering / linear diffusion - the most widely used method

- mid 80’s – unified formulations – a breakthrough!
  - methods that combine smoothing and edge detection (Geman & Geman’84, Mumford & Shah’89, Perona & Malik’90)
A little bit of (personal) history

Standard unified formulations (nonlinear filters) fail to capture some details, e.g. due to texture!

- mid 80's – unified formulations – a breakthrough!
  - methods that combine smoothing and edge detection (Geman & Geman'84, Mumford & Shah'89, Perona & Malik'90)
Some seminal works

- **Bilateral Filter**
  - Tomasi and Manduchi 1998
  - Durand and Dorsey 2002

- **Total Variation**
  - Rudin et al. 1992

- **Fast Cartoon + Texture**
  - Buades et al. 2010

- **Relative Total Variation**
  - Xu et al. 2012

- **L0 Smoothing**
  - Xu et al. 2011

- **Relative Total Variation**
  - Xu et al. 2012

- **Envelope Extraction**
  - Subr et al. 2009

- **Context-guided Filtering**
  - Erdem and Tari 2009

- **RegCov Smoothing**
  - Karacan et al. 2013

### Figures

- **Fig. 10** Coalition of three feedback measures.
  - (a) Source image.
  - (b) Reconstruction result with contextual feedback.
  - (c) Final edge indicator function.
  - (d) Texture edges measure $\phi_{te}$.
  - (e) Reconstruction using Shah's modification.
  - (f) Reconstruction using the ROF model.

- **Fig. 11** Texture preserving denoising with local scale measure $\phi_{ls}$.
  - (a) Source image.
  - (b)–(c) Reconstructions using AT with different choices of smoothing levels.
  - (d) Reconstruction result with contextual feedback. Notice that textures in the fabric are preserved.
  - (e) Local scale measure $\phi_{ls}$.

Increasing the level of smoothing in the AT model results in noise-free results as presented in Fig. 11(c), however the textured regions are also smoothed out during the process. Figure 12 illustrates the results of two more texture preserving denoising experiments. Figures 12(c) and (d) are obtained using the parameters $\alpha = 20$, $\beta = 0.1$, $\rho = 0.001$, $\epsilon_{ls} = 0.125$, $n = 15$ and $\alpha = 4$, $\beta = 0.1$.
Some seminal works

Bilateral Filter
Tomasi and Manduchi 1998
Durand and Dorsey 2002

Total Variation
Rudin et al. 1992

Fast Cartoon + Texture
Buades et al. 2010

Relative Total Variation
Xu et al. 2012

WLS Filter
Farbman et al. 2008

L0 Smoothing
Xu et al. 2011

Envelope Extraction
Subr et al. 2009

Bilateral Filter
Tomasi and Manduchi 1998
Durand and Dorsey 2002

WLS Filter
Farbman et al. 2008

Envelope Extraction
Subr et al. 2009

Context-guided Filtering
Erdem and Tari 2009

RegCov Smoothing
Karacan et al. 2013

RegCov Smoothing
Karacan et al. 2013
Context-guided filtering

- Contextual knowledge extracted from local image regions guides the regularization process.

Image Credit: P. Milanfar
Structure-Texture Decomposition

- Decomposing an image into structure and texture components
Structure-Texture Decomposition

- Decomposing an image into structure and texture components
Structure-Texture Decomposition

- Decomposing an image into structure and texture components
Structure-Texture Decomposition

- Decomposing an image into structure and texture components
Region Covariances as Region Descriptors

Motivated by these properties, in this study we employ the region and local structures are described by similar covariance matrices.

Expressing an image region by the covariance of features extracted at pixels (Equation 6).

A naive implementation of our structure preserving image smoothing approximation of the Mahalanobis distance between two Normal distributions with

\[ \mathbf{C}_R = \frac{1}{n-1} \sum_{i=0}^{n} (\mathbf{z}_k - \mu)(\mathbf{z}_k - \mu)^T \]

In the experiments, we handle color images by computing the patch correlation of other features with the spatial coordinates. The features, namely intensity, orientation, and pixel coordinates so that an image pixel is represented with a 7-dimensional feature vector:

\[ F^C(p) = \left[ \begin{array}{c} I(p) \frac{\partial I}{\partial x} \frac{\partial^2 I}{\partial x^2} \frac{\partial^2 I}{\partial y^2} \end{array} \right] \]

The key to our adaptive filtering framework relies on how we decompose the covariance matrix (symmetric positive semi-definite matrix) has a unique Cholesky decomposition and use it to transform covariance differences in means. Therefore, in this paper, we investigated the relationship in means and covariances of features extracted from the image patches, here we propose two alternative schemes based on the following parameterisation:

\[ F(x, y) = \phi(I, x, y) \]

where \( \phi \) can be a squared neighborhood centered at \( (x, y) \) using either Eq. 8 (Model 1) or

\[ \mathbf{C}_R = \frac{1}{n-1} \sum_{i=0}^{n} (\mathbf{z}_k - \mu)(\mathbf{z}_k - \mu)^T \]
Motivation

- Region covariances capture local structure and texture information.
- Similar regions have similar statistics.
RegCov Smoothing - Formulation

\[ I = S + T \]

\[ S(p) = \frac{1}{Z_p} \sum_{q \in N(p, r)} w_{pq} I(q) \]

- Structure-texture decomposition via smoothing
- Smoothing as weighted averaging
- Different kernels \( w_{pq} \) result in different types of filters.
- Three novel patch-based kernels for structure texture decomposition.

Model 1

- Depends on sigma-points representation of covariance matrices (Hong et al., CVPR’09)

\[
\mathbf{C} = \mathbf{L} \mathbf{L}^T \quad \text{Cholesky Decomposition} \quad S = \{s_i\} \quad \text{Sigma Points}
\]

\[
s_i = \begin{cases} 
\alpha \sqrt{d} \mathbf{L}_i & \text{if } 1 \leq i \leq d \\
-\alpha \sqrt{d} \mathbf{L}_i & \text{if } d + 1 \leq i \leq 2d 
\end{cases}
\]

**Final representation**

\[
\Psi(\mathbf{C}) = (\mu, s_1, \ldots, s_d, s_{d+1}, \ldots, s_{2d})^T
\]

**Resulting kernel function**

\[
w_{pq} \propto \exp \left( - \frac{\|\Psi(\mathbf{C}_p) - \Psi(\mathbf{C}_q)\|^2}{2\sigma^2} \right)
\]
Model 2

- An alternative way is to use statistical similarity measures.
- A Mahalanobis-like distance measure to compare to image patches.

\[
d(p, q) = \sqrt{(\mu_p - \mu_q)C^{-1}(\mu_p - \mu_q)^T}
\]

\[
C = C_p + C_q
\]

*Resulting kernel* \( w_{pq} \propto \exp\left(-\frac{d(p, q)^2}{2\sigma^2}\right) \)
Model 3
resulted from a discussion with Rahul Narain (Berkeley University)

- We use Kullback-Leibler (KL)-Divergence measure from probability theory.
- A KL-Divergence form is used to calculate statistical distance between two multivariate normal distributions.

\[
d_{KL}(p, q) = \frac{1}{2} \left( tr(C_q^{-1}C_p) + (\mu_p - \mu_q)^T C_q^{-1}(\mu_p - \mu_q) - k - \ln \left( \frac{\det C_p}{\det C_q} \right) \right)
\]

**Resulting kernel**

\[
w_{pq} \propto \frac{d_{KL}(p, q)}{2\sigma^2}
\]
Smoothing Kernels

Figure 3.2.: Our filtering kernels consider local image geometry on calculation of filtering weights by capturing texture information.

\[121\] and \((x, y)\) denotes the pixel location. Hence, the covariance descriptor of an image patch is computed as a 7 \(\times\) 7 matrix. Including \((x, y)\) into the feature set is important since it allows us to encode the correlation of other features with the spatial coordinates. The feature set can be extended to include other features, like for example rotationally invariant forms of the derivatives, if desired.

In the experiments, we handle color images by computing the patch similarity weights \(w_{pq}\) using the intensity information and taking the weighted average over the corresponding RGB vectors rather than the intensity values in Equation 23. We empirically found that including RGB components to the feature set does not change the results much but increases the running times.

3.3. Model 1

Using the set \(S\) defined by Equation (21), a vectorial representation of a covariance matrix can be obtained by simply concatenating the elements of \(S\). Moreover, first-order statistics can be easily incorporated to this representation scheme by including the mean vector of the...
Input
Model2 Texture
TV
Rudin et al.
1992
Envelope Extraction
Subr et al.
2009
Model 1
Model 2
Model 3
• Shading preserved

• Structure preserved

• No unintuitive edge
Input

WLS

Local Extrema

RTV

Preserves shade and structure

Model2
Multiscale decomposition
$S_1(k = 5)$
$S_2(k = 7)$
Model2+Model1

Input

Model2

Model2 Texture

Model2+Model1
Edge Detection
Canny edges of original image

Canny edges of smoothed image
Image Abstraction
Detail Boosting
Image Composition
Inverse Halftoning
Image Retargeting
Extracted Seams
Avidan and Shamir 2007
Retargeting Results
Avidan and Shamir 2007
Where we are going

- Linear filtering
- Nonlinear filtering (unified formulations)
- Pixels to Patches (context is more important than content)
- New patch representations may reveal new smoothing behaviors
- Better the smoothing, better the applications!
- Clearly, we have a long way to go to solve the problem of image smoothing!