Subjective continuity: combining perceptual and metric extrapolation

Shay Gepstien and Yosi Keller*

School of Engineering, Bar-Ilan University

Outline

1. Quick review of image inpainting
   - PDE based approach
   - Sampling based approaches
2. Quick review of spectral embedding
3. Image inpainting in the diffusion domain

Image inpainting

M. Bertalmeo, G. Sapiro, V. Caselles and C. Ballester.
SIGGRAPH 2000

Heat equation I

Basic idea (over simplified)

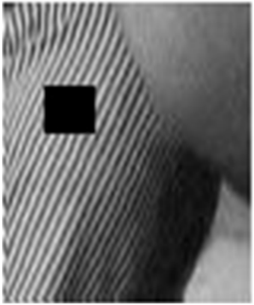
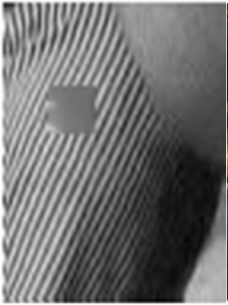
$$I^* = \min_I \iint |\nabla I|^2 \, ds, \; I(\Omega) = I_0$$

by calculus of variations we get the heat equation

$$\triangle I = 0, \; I(\Omega) = I_0$$

- Bertalmeo *et al.* used a non-linear heat equation
- Information is propagated outward-in
- Restores smooth image regions and cracks
- Has difficulties handling texture and large holes

Heat equation II
Basic idea (over simplified)


Extensions:

1. Anisotropic Diffusion
2. The U+V model: Bertalmio, Sapiro, Vese, Osher, TIP2003

Heat equation III

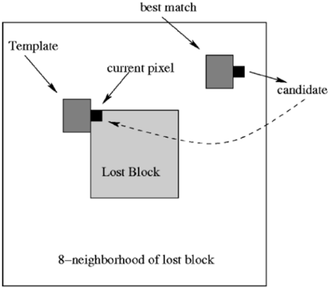
Basic idea (over simplified)

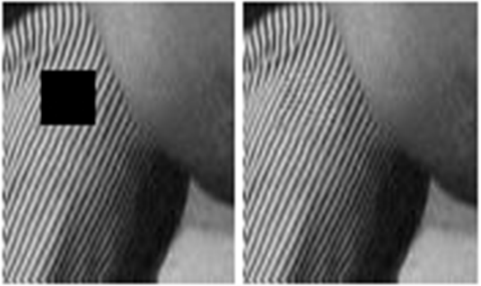
The good:
1. a clear objective function to optimize
2. assumes a simplistic manifold structure

The bad:
1. the numerical implementation can only recover smooth functions
2. The algorithm only learns from the boundary (implicit downside)

Patch sampling I

Introduced in the seminal work of Efros and Leung CVPR1999



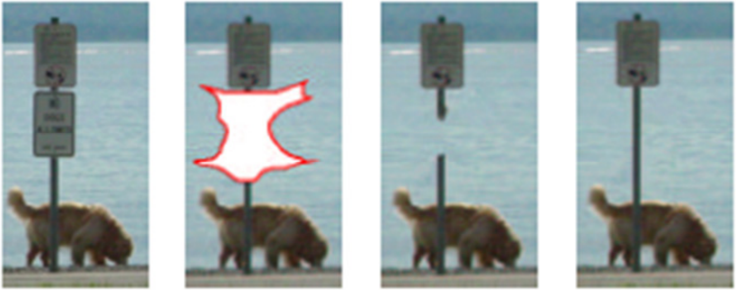


Patch sampling II

Basic idea (in signal processing terms)

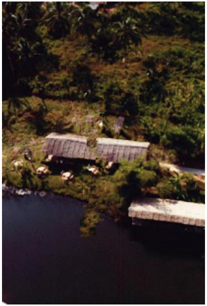
The analogue of oversampling an analog signal in order to reconstruct it using a ZOH filter.

Extended by Criminisi, Perez, and Toyama CVPR2003



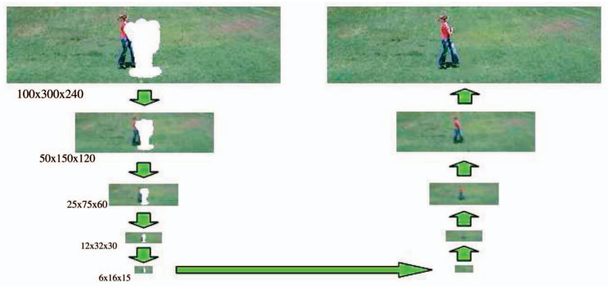
Patch sampling III

- An anisotropic formulation of Efros' algorithm: propagates patches along edges
- Automatically maintain the structure during synthesis



Further extension by Michal Irani 2004: multiscale and iterations

Patch sampling IV



The good:
1. Able to inpaint large and textured regions
2. Learns from the entire image, can also learn from multiple images
3. Very good results

The bad:
1. No objective function to optimize

Patch sampling V

2. Computationally exhaustive
3. Does not utilize the manifold structure

Kernel methods

Definition

Given a dataset $\{x_i\}_{i=1..n}$:

1. Apply a p.s.d. kernel k to $\{x_i\}$. For instance: $w_{ij} = \exp(-\frac{d(x_i,x_j)}{\varepsilon})$, $\varepsilon > 0$.
2. Compute the eigenvectors of W: $w_{ij} = \sum_{l \geq 0} \lambda_l \psi_l(i) \psi_l(j)$,
   The embedding is given by
3. $\Psi(x_i) : x_i \mapsto \left(\lambda_1 \psi_1(x_i), \lambda_2 \psi_2(x_i), \ldots\right)$
4. $\|\Psi_t(x_i) - \Psi_t(z_i)\|^2_{L_2} = \sum_{l=0}^{n-1} \lambda_l^{2t} (\psi_l(x) - \psi_l(z))^2 =$
   $w_{ii} + w_{jj} - 2w_{ij} = D(x,z)^2$

Different views of spectral embedding

- Spectral graph partitioning: Shi and Malik *Normalized Cuts and Image Segmentation*, 98
- Random Walk interpretation: Meila and Shi, *A random walks view of spectral segmentation*
- Preserving the infinitesimal geometry, Laplacian eigenmaps, Belkin-Nyogi, NIPS2002
- Diffusion distances, Coifman et. al. PNAS2004
  $$\|\Psi(x_i) - \Psi(z_i)\|_{L_2}^2 = \sum_{y \in \Omega} \frac{(p(x,y) - p(z,y))^2}{\phi_0(y)}$$
- Asymptotic approach: Diffusion vectors as an approximation of the eigenvectors of the Fokker-Planck operators. S. Lafon and B. Nadler. NIPS2005.

Signal processing in diffusion space
Joint work with Shay Gepstien

- Is this is an interesting image?

Signal processing in diffusion space
Joint work with Shay Gepstien

- Is this is an interesting image? No
- Is it easy to inpaint?

Signal processing in diffusion space
Joint work with Shay Gepstien

- Is this is an interesting image? No
- Is it easy to inpaint? Yes
- Is it easy to compress?

Signal processing in diffusion space

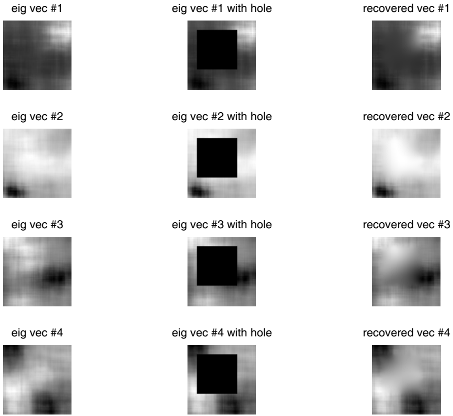
Joint work with Shay Gepstien

- Is this is an interesting image? No
- Is it easy to inpaint? Yes
- Is it easy to compress? No

Looks (metrics) are deceiving!

Constructing our own metrics and embedding (and bases)

For an appropriate metric the embeddings are smooth



LBP texture features + Heat equation

Is the embedding reversible?

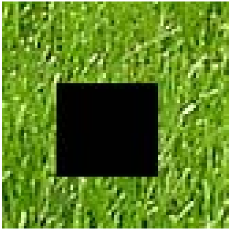
Is the embedding reversible?

The short answer is: no

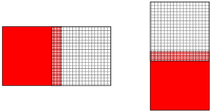
- The embedding is a non-linear operation that is non-invertible.
- The embedding captures the intrinsic structure of a signal and not its ambient space manifestation.

The long answer is: we can try

- For each *interpolated* point $\{\phi_i(j)\}_{i=1}^{N}$ we can pick the $K-NN$ points in the learning set
- We can impose (constrained) spatial smoothness
  - PDE - smoothness over greyscale values
  - Patches - smoothness over *all* patches



We define the pairwise distance $d\left(x_i, x_b\right)$ using the L_2 norm



$$W\left(x_i, x_j\right) = \exp -\frac{d^2\left(x_i, x_j\right)}{\sigma^2}$$

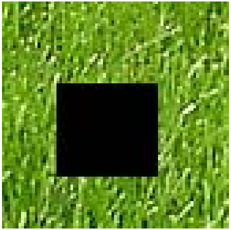
and maximize

$$\mathbf{x}^* = \max \arg_{\mathbf{x}} \sum_{i,j} W\left(x_i, x_j\right)$$

Similar to: S. Avidan *et al.* : Seam Carving, Image Retargeting
Main difference: we only consider K patches per point

Initial results

Source



Result

Source

Result

Comparison to previous works

The good:

1. a globally optimal solution in high dimensional space
2. the diffusion is applied to low dimensional *smooth* data
3. can be applied to different data sources

The bad: there are a few engineering issues to resolve

Conclusions

- Global optimality goes hand in hand with smoothness
- Image patches can be used as building blocks for image inpainting
- Image patches are part of a manifold: a fact not exploited by current schemes
- Image inpainting is an example of a class of general problems: how to utilize manifolds in signal processing
- Future work: can we apply *other* signal processing operations and “invert” the embedding?
  - Filtering
  - Compression
- Can we apply it to signals other than images (speech)?

Thanks You!