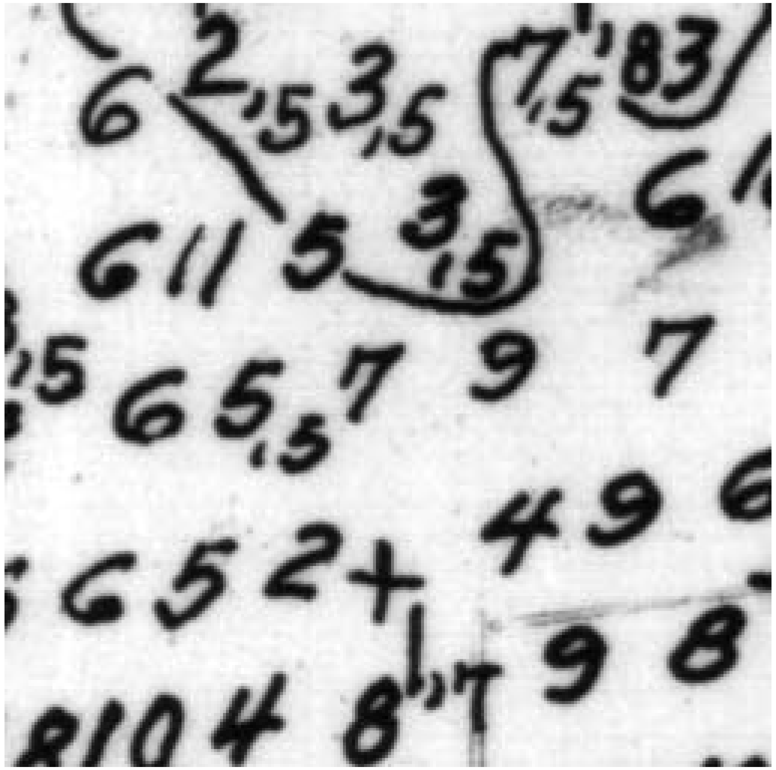
An Algebraic Multigrid Approach for Shape from Photometric Stereo and Binarization

Ron Kimmel and Irad Yavneh
Department of Computer Science
Technion

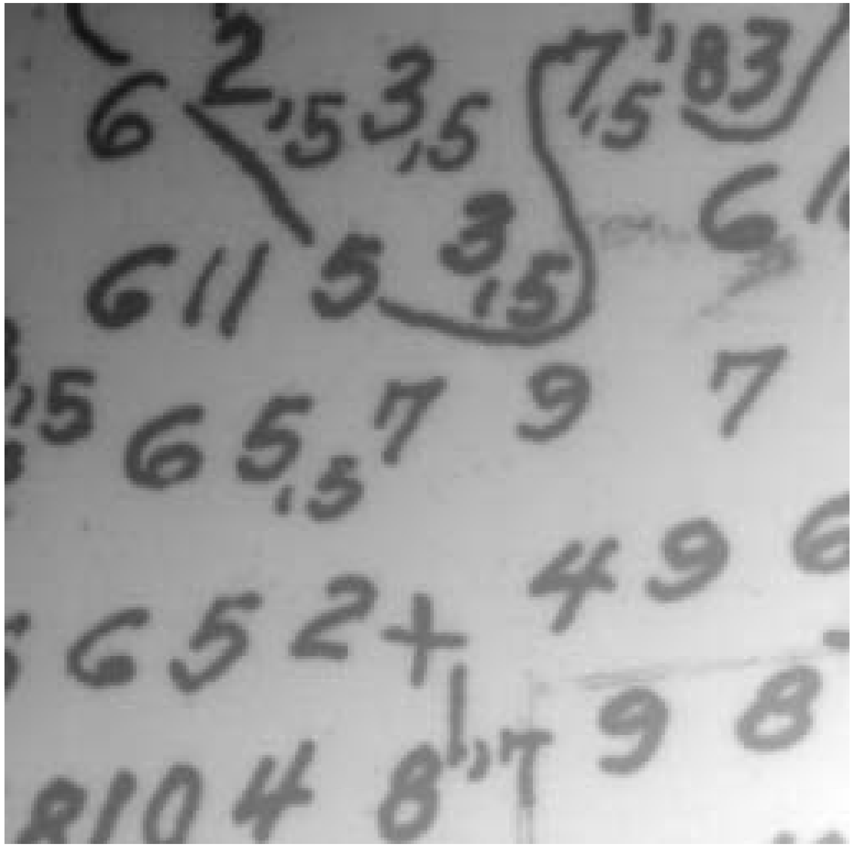
Based on: [Kimmel and Y, *SISC* **24** (4), p. 1218, 2003]

Image Binarization

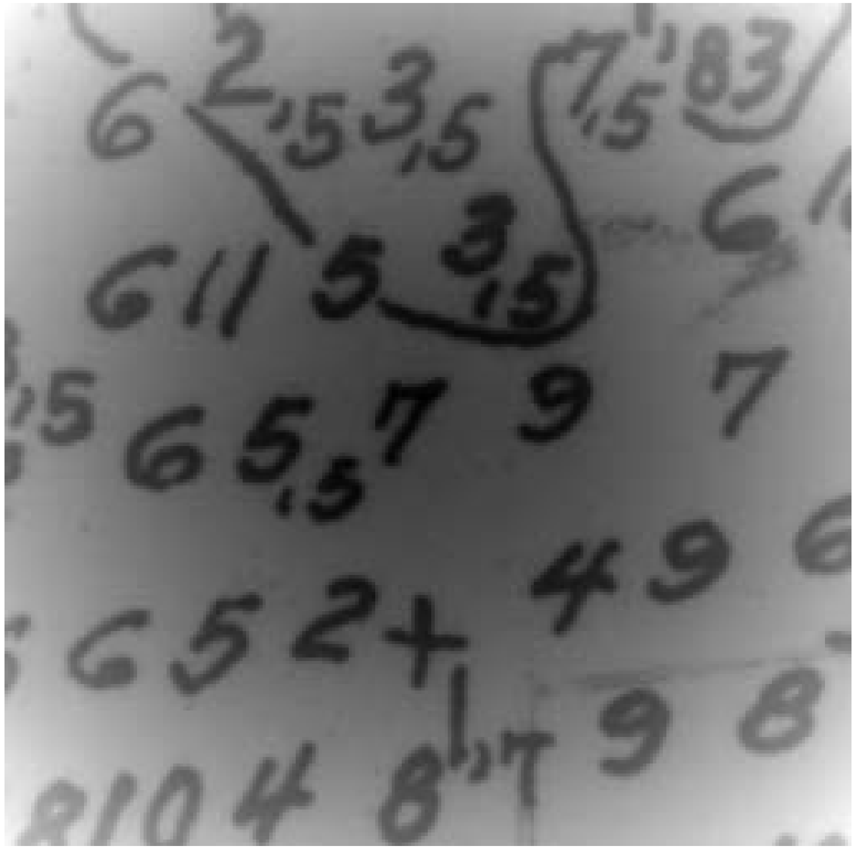


Original Image

Nonuniform Illumination

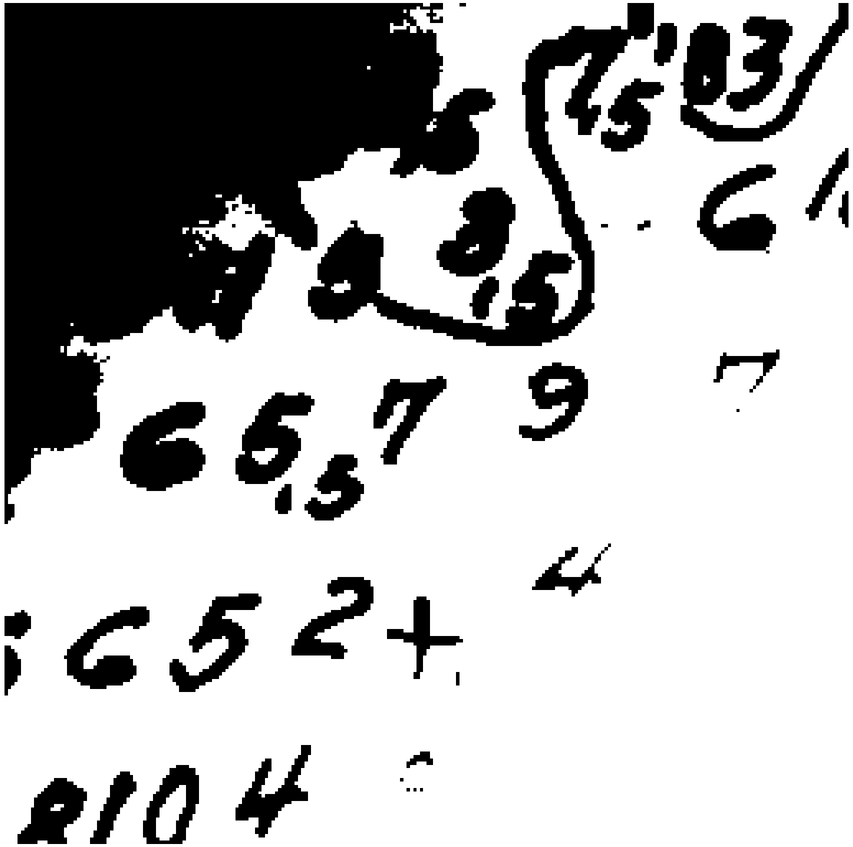


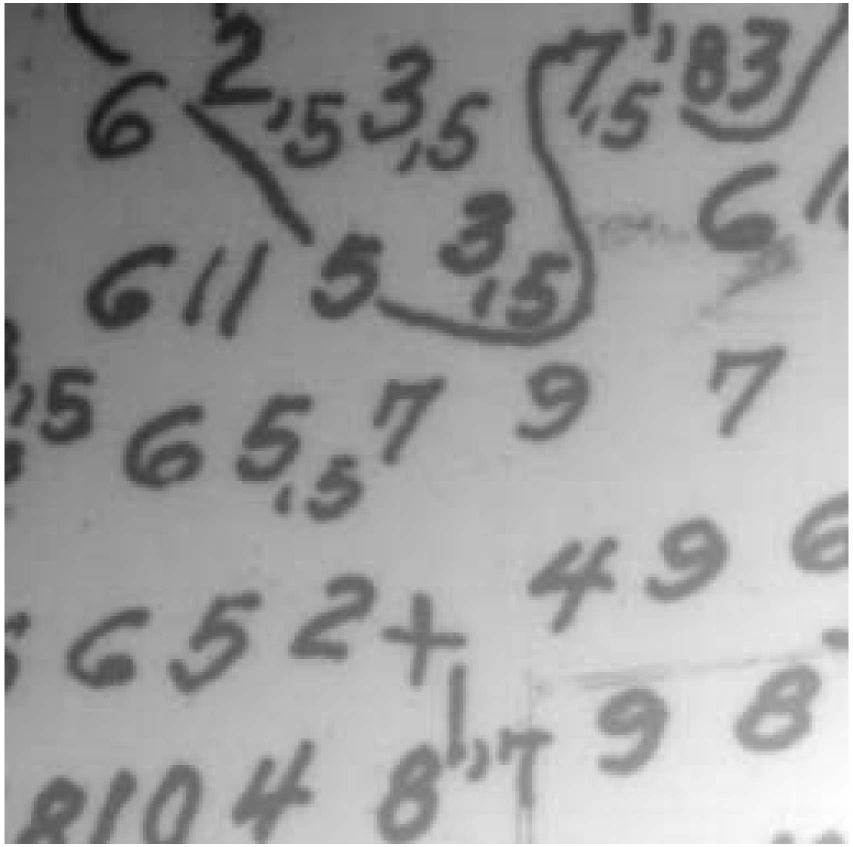
Tilted



Spherical

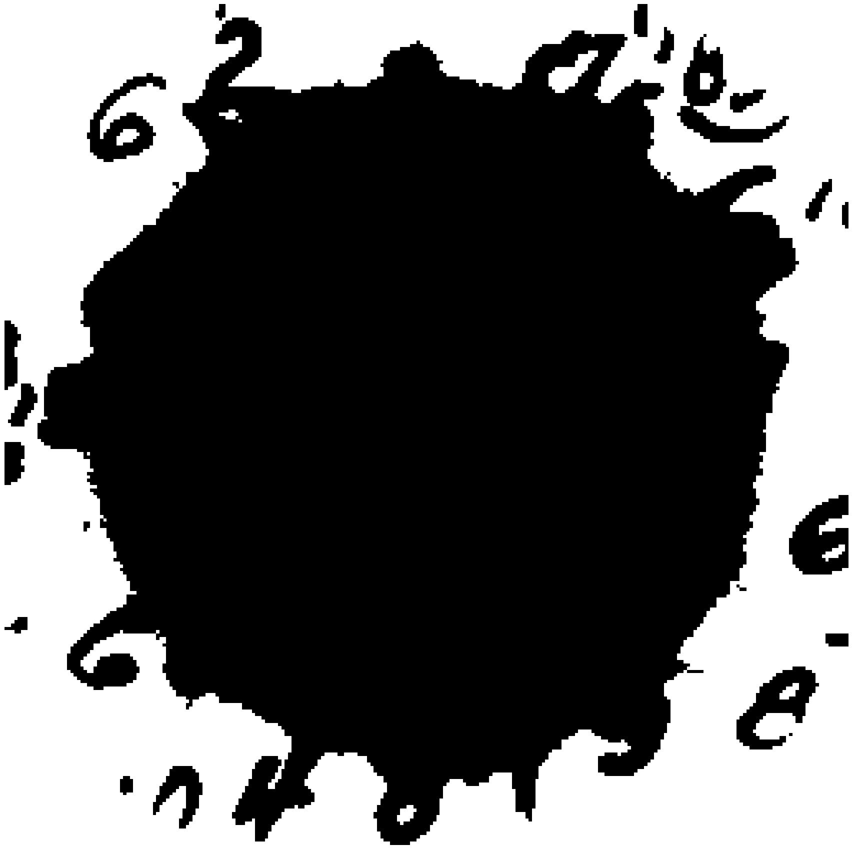
Naïve (threshold) binarization

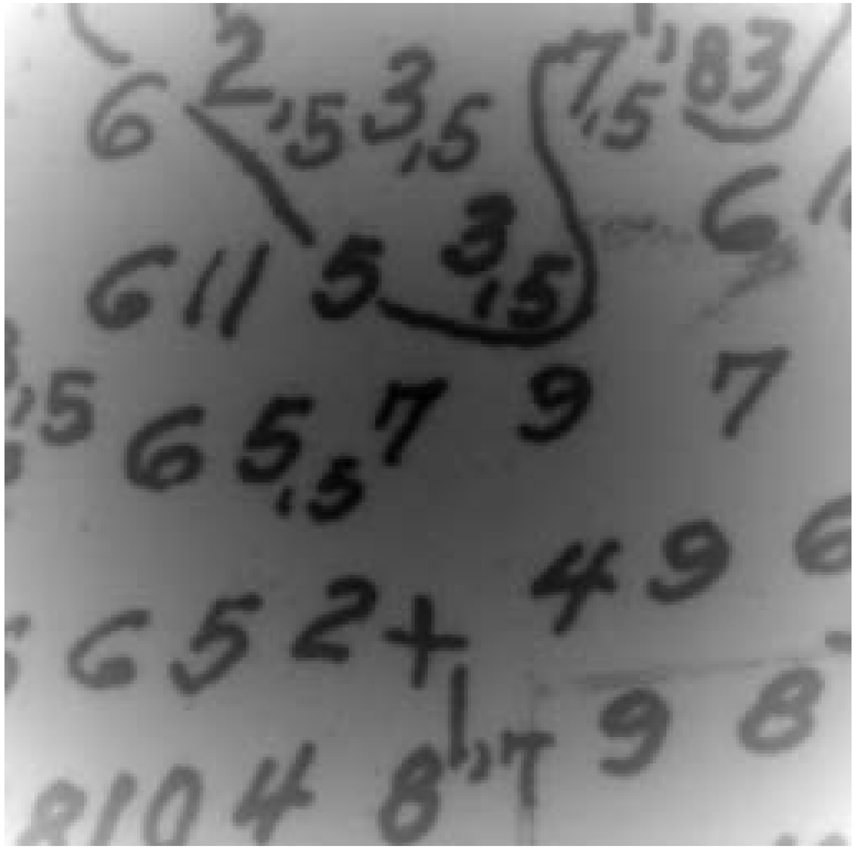




Tilted

Naïve (threshold) binarization





Spherical

Yanowitz-Bruckstein Binarization

- Isolate the locations of edge centers, for example, the set of points,

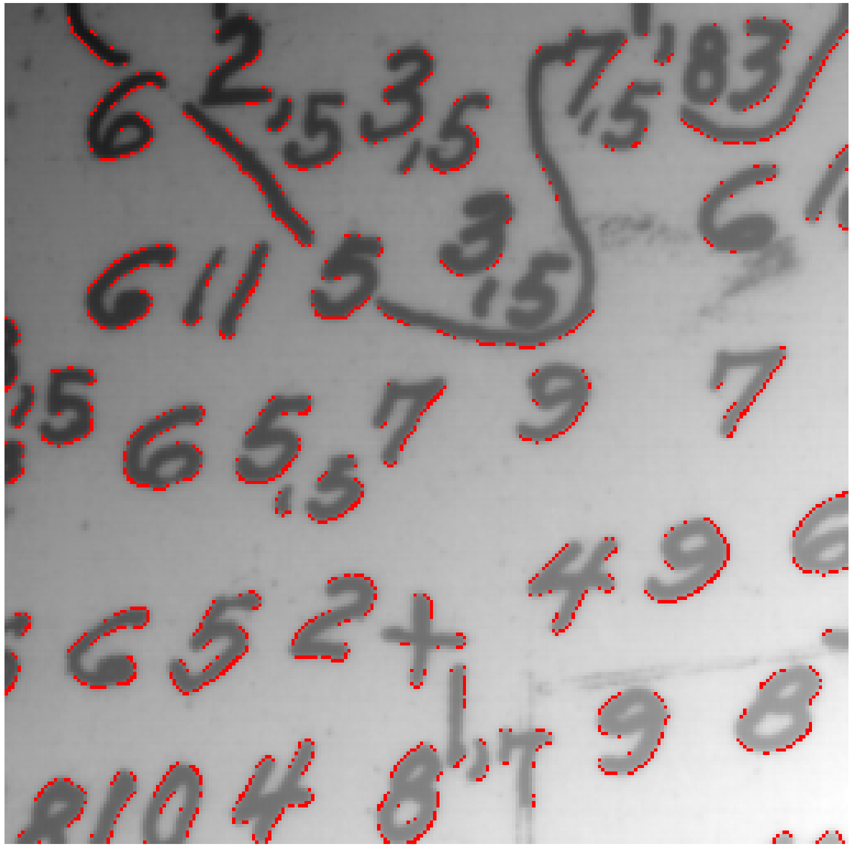
$$s = \{(x, y) : |\nabla I| > T\}$$

  for some threshold T.

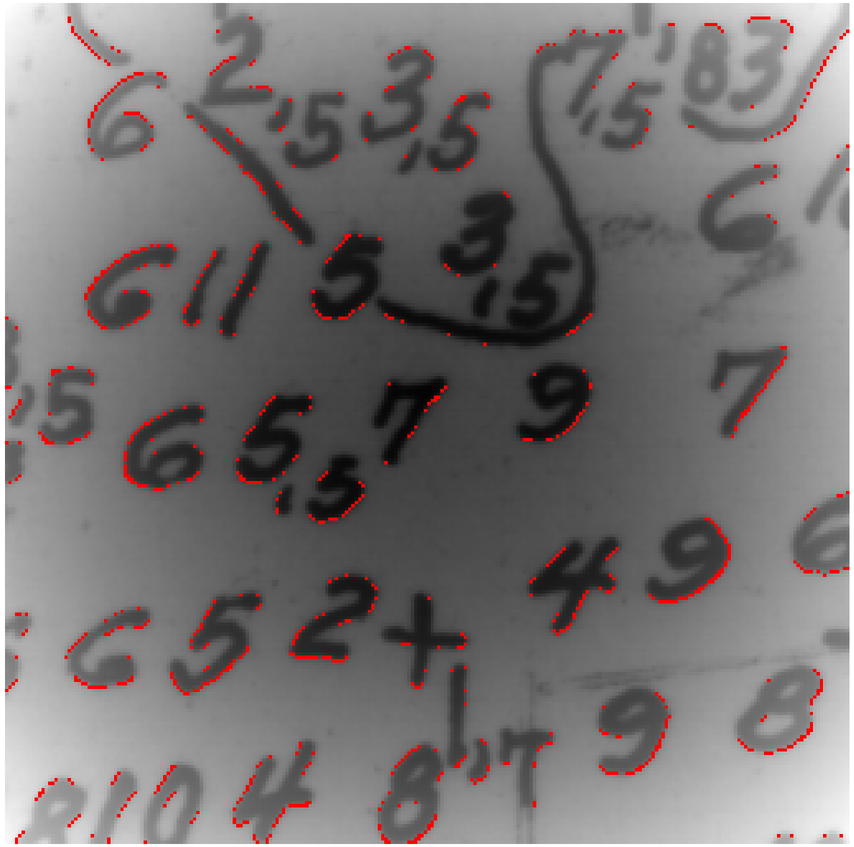
- Use the values $I(x,y)$, for (x,y) in s, as constraints for a threshold surface, u, which elsewhere satisfies the equation

$$\Delta u = 0.$$

Edges

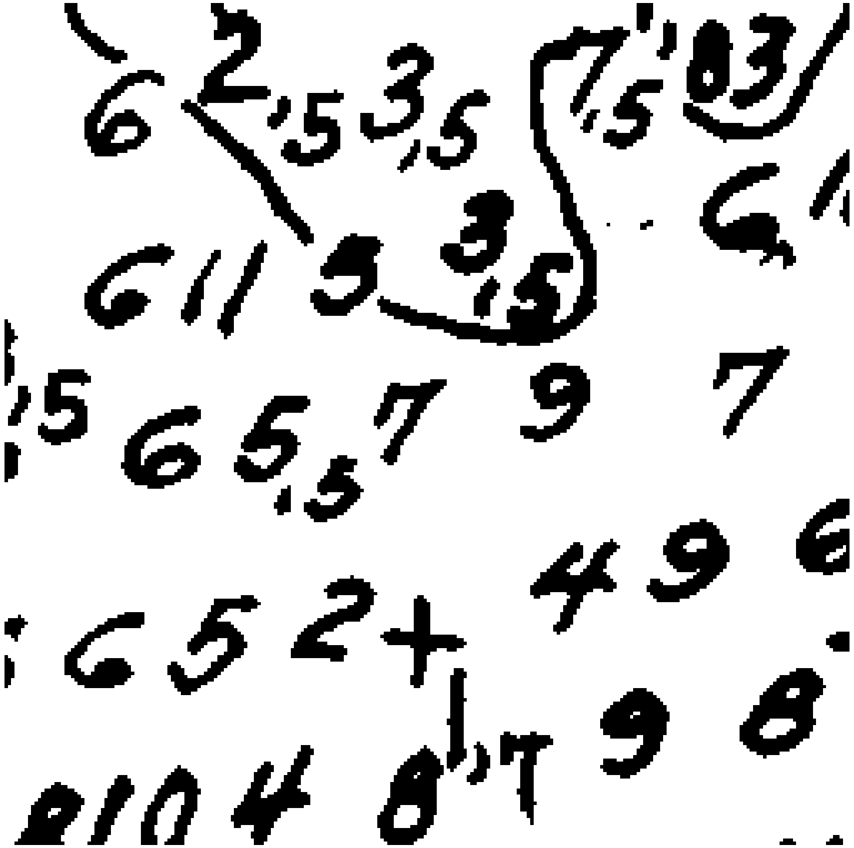


Tilted

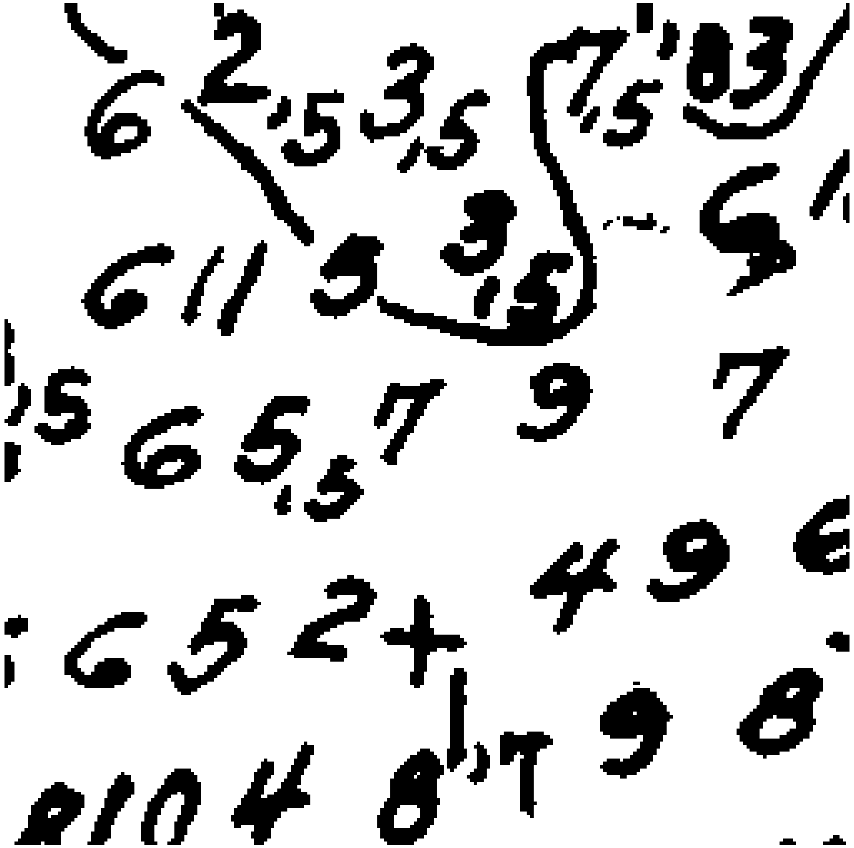


Spherical

Results



Tilted



Spherical

Algebraic Multigrid for Shape from Photometric Stereo

1. What is “shape from photometric stereo”?
2. Why is the “standard approach” insufficient?
3. What can we do about it?
4. How does (algebraic) multigrid help?
5. Some results

The Problem

"Shape from photometric stereo" deals with the problem of shape reconstruction from 2D projections of the real world onto a camera in the case where the camera is fixed and several images are obtained with different lighting conditions.

Input images with the same camera position and head object but three different lighting directions.

Lambertian reflectance model: Given three images, I_1, I_2, I_3, of the same object with three different lighting directions, l_1, l_2, l_3, we assume

$$I_i = \rho\langle l_i, N\rangle, \quad i = 1,2,3,$$

where $\rho(x, y)$ is the albedo (which depends on the properties of the object), and N is the normal to the surface, $z(x,y)$, given by

$$N = \frac{(-z_x, -z_y, 1)}{\sqrt{1 + z_x^2 + z_y^2}}.$$

The approximate surface gradient,

$$(p, q)^T \approx \nabla z$$

can be extracted easily from the images.

We wish to reconstruct $z(x,y)$, keeping in mind that there are errors in the model and in the measurements.

Variational formulation: find the surface which minimizes the functional

$$\iint_\Omega w(x,y)\left\|(p,q)^T - \nabla z\right\|^2 dx\,dy,$$

with $w(x,y) > 0$. The Euler-Lagrange equation is

$$\nabla \cdot \left(w(x,y)\left(p - z_x, q - z_y\right)^T\right) = 0, \quad (x,y) \in \Omega,$$

$$\nabla z \cdot \mathbf{n} = (p,q)^T \cdot \mathbf{n}, \quad (x,y) \in \partial\Omega,$$

where $\mathbf{n}$ is the outwards normal to $\partial\Omega$.

The choice $w(x, y) \equiv 1$ yields the Poisson equation. The resulting reconstruction is often unsatisfactory due to errors in the model and in the measurements, shaded regions, etc.

Additional data: constraints at points where the height z is known accurately by some independent measurement.

We then minimize the functional subject to the c constraints,

$$z\left(x_k, y_k\right) = z_k, \quad k = 1, \ldots, c.$$

Now, however, $w(x,y)=1$ will result in spurious "spikes" at the constrained points. Indeed, one can show that in order to maintain p continuous derivatives of the reconstructed shape at $(x, y) = (x_k, y_k)$ we must have

$$w(x, y) \approx r_k^{-\alpha},$$

with $\alpha > p$ in a neighborhood of (x_k, y_k), where

$$r_k(x, y) = \sqrt{(x - x_k)^2 + (y - y_k)^2}.$$

Accordingly, we set

$$w(x, y) = \left(\min_k r_k\right)^{-2}.$$

There exists an efficient algorithm to compute $w(x,y)$.

Computationally, the problem is now more difficult due to the constraints and the singular w.

Approaches based on simple multigrid or FFT are inefficient.

Hence, we choose a novel robust algebraic multigrid approach:

- Galerkin coarsening
- Specialized Prolongations

Some Results: Side and perspective views of the reconstructed surface, with the frontal textured mapped onto it. Left: unconstrained; Right: eight constrained points.

A multigrid approach for multidimensional scaling (MDS)

Michael M. Bronstein, Alexander Bronstein
Ron Kimmel, Irad Yavneh

Department of Computer Science

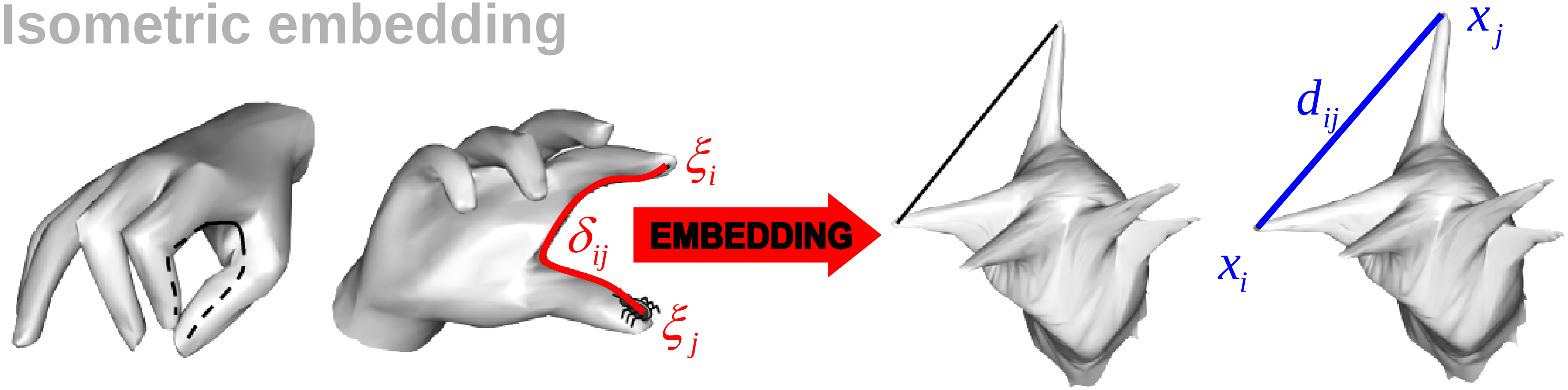
Technion – Israel Institute of Technology

Based on: [B^2,K,Y: *NLAA* 13(2-3), 149-171, 2006]

Plan

- Isometric embedding problem
- Motivation from 3D face recognition
- Basic MDS algorithm
- Multigrid MDS
- Results

Isometric embedding



A mapping between finite metric spaces

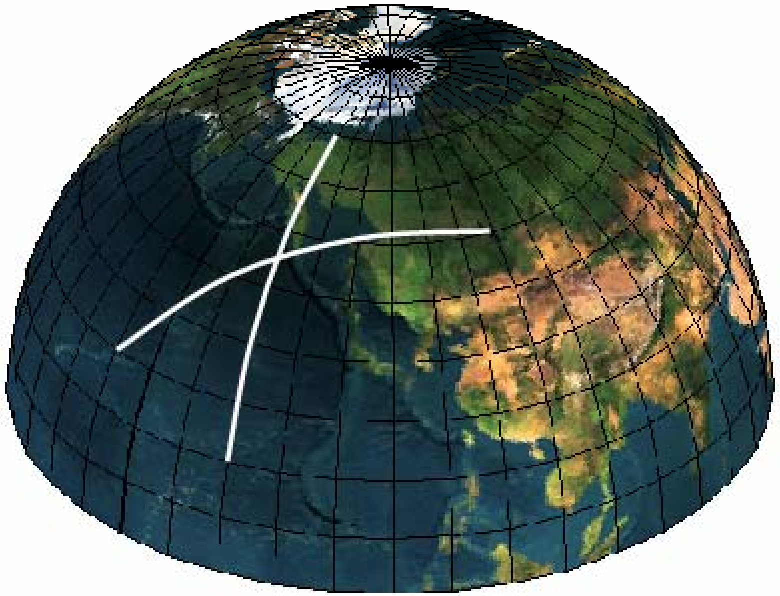
$$\varphi : \left(\{ \xi_1, \ldots, \xi_N \} \subset \mathcal{S}, \boldsymbol{\Delta} \right) \to \left(\{ x_1, \ldots, x_N \} \subset \mathbb{R}^m, \mathbf{D} \right)$$

such that $\delta_{ij} \approx d_{ij} = \left\| x_i - x_j \right\| \quad \forall i, j = 1, \ldots, N.$

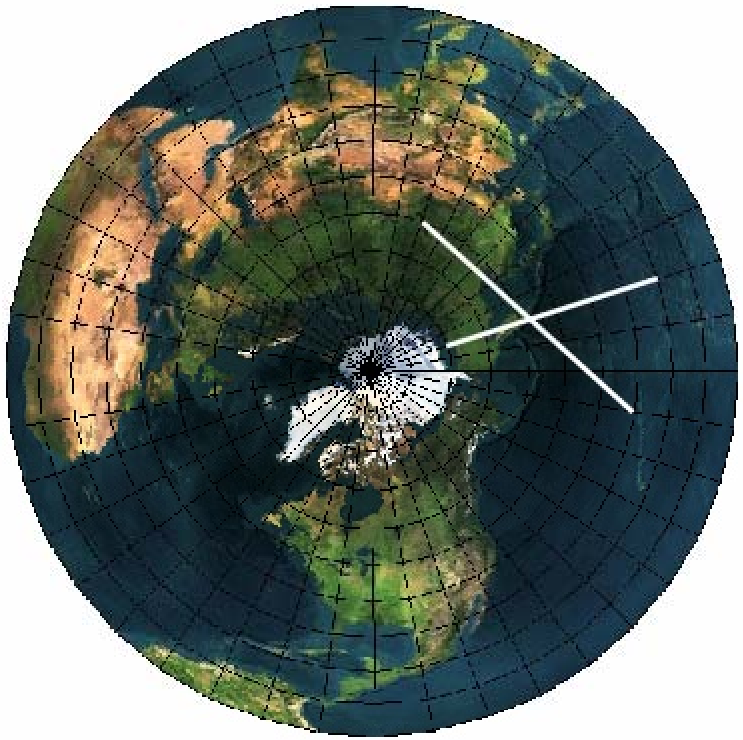
- $\mathbb{R}^m$ - m-dimensional **embedding space**
- $\boldsymbol{\Delta} = \left(\delta_{ij} \right)$ - $N \times N$ matrix of original **geodesic distances**
- $\mathbf{D} = \left(d_{ij} \right)$ - $N \times N$ matrix of distances in the embedding space

A. Elad, R. Kimmel, CVPR 2001

Isometric embedding in cartography



GLOBE (HEMISPHERE)



PLANAR MAP

Exact isometric embedding of the sphere into any $\mathbf{R}^m$ does not exist

Multidimensional scaling (MDS)

- **Exact** isometric embedding does not exist in most cases
- MDS = minimization of **embedding error** criterion (**stress**), e.g. LS:

$$s(\mathbf{X}) \equiv \sum_{i<j} \left(d_{ij}(\mathbf{X}) - \delta_{ij} \right)^2$$

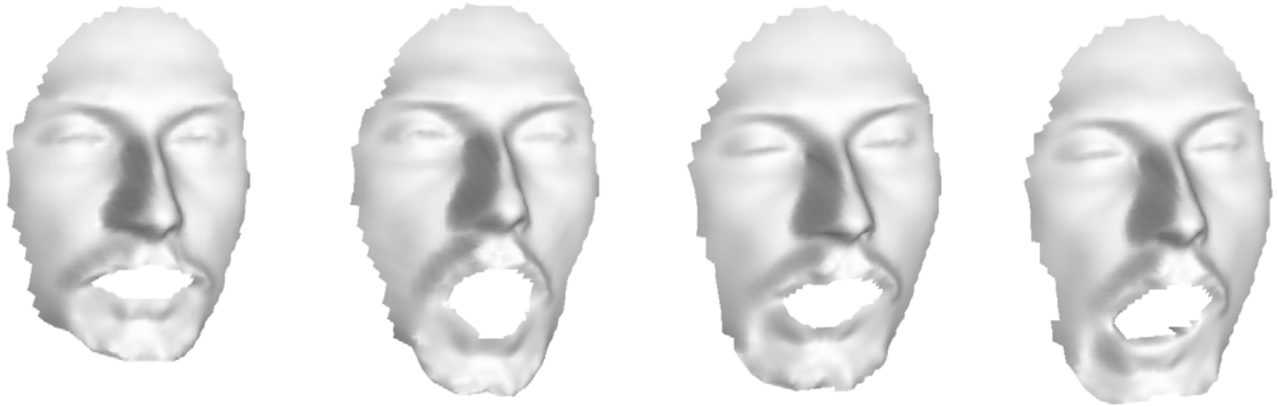
- $\mathbf{X} = (\mathbf{x}_1; \ldots; \mathbf{x}_N)$ - $N \times m$ matrix of coordinates in $\mathbb{R}^m$
- Nm optimization variables
- Optimum defined up to an **isometry group** in $\mathbb{R}^m$

I. Borg, P. Grönen, *Modern multidimensional scaling*, Springer, 1997

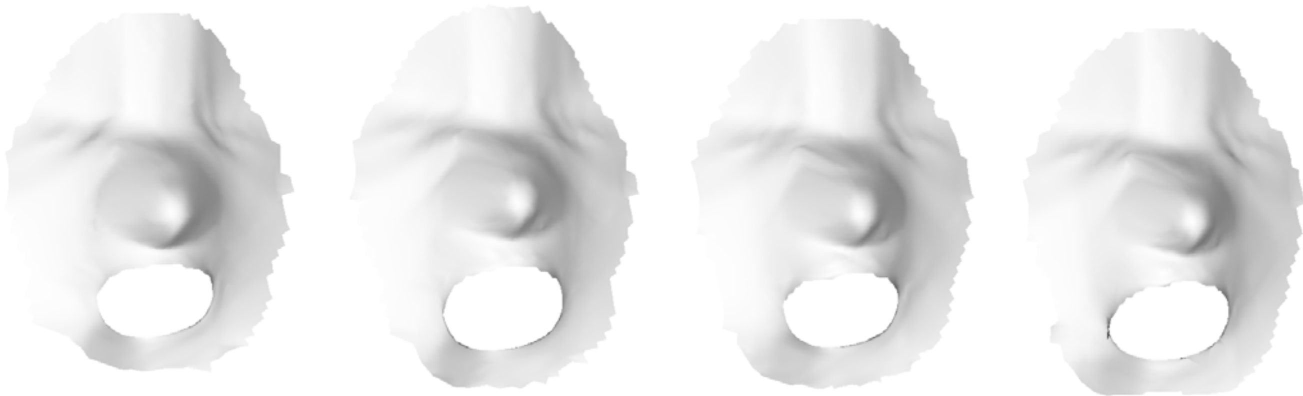
Caveats

- **Non-convex** and **nonlinear** optimization problem
- Hessian is structured but **full (dense)**
- Computational complexity of $s(\mathbf{X})$ and $\nabla s(\mathbf{X})$ is approximately the same
- Newton algorithm is prohibitive for large N
- Line search is disadvantageous

Isometric embedding in 3D face recognition



FACIAL SURFACES

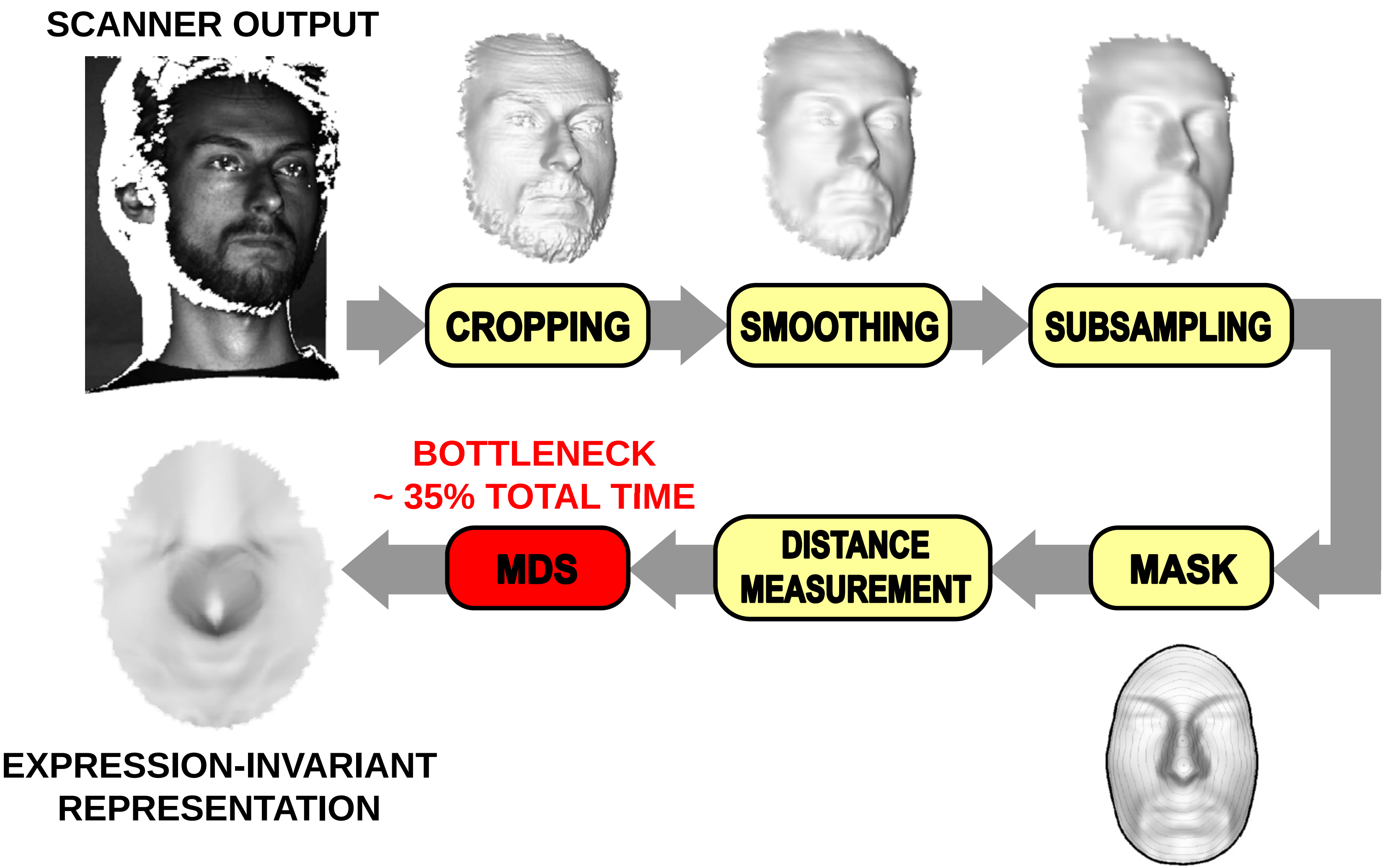


EXPRESSION-INVARIANT REPRESENTATIONS (EMBEDDING INTO $\mathbb{R}^3$)

M. Bronstein, A. Bronstein, R. Kimmel, “Expression-invariant representation for human faces”

3DFACE face recognition system

3DFACE face recognition system



SMACOF – a basic MDS algorithm

$$\min_{\mathbf{X}} \sum_{i<j} \left(d_{ij}(\mathbf{X}) - \delta_{ij} \right)^2$$

- First-order gradient-descent type optimization method
- Gradient of the stress function:

$$\nabla_{\mathbf{X}} s(\mathbf{X}) = 2\mathbf{U}\mathbf{X} - 2\mathrm{B}(\mathbf{X})\mathbf{X}$$

where

$$u_{ij} = \begin{cases} -1 & \text{if } i \neq j \\ N-1 & \text{if } i = j \end{cases} \qquad b_{ij} = \begin{cases} -\delta_{ij} d_{ij}^{-1}(\mathbf{X}) & \text{if } i \neq j \text{ and } d_{ij} \neq 0 \\ 0 & \text{if } i \neq j \text{ and } d_{ij} = 0 \\ -\sum_{j \neq i} b_{ij} & \text{if } i = j \end{cases}$$

- A gradient descent step can be performed with a multiplicative update

$$\boxed{\mathbf{X}^{(k+1)} = \mathbf{U}^{\dagger}\mathrm{B}\left(\mathbf{X}^{(k)}\right)\mathbf{X}^{(k)}} \quad \Leftrightarrow \quad \mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} - \frac{1}{2}\mathbf{U}^{\dagger}\nabla_{\mathbf{X}} s\left(\mathbf{X}^{(k)}\right)$$

SMACOF STEP

I. Borg, P. Grönen, *Modern multidimensional scaling*, Springer, 1997

Multigrid MDS components

- **Hierarchy of grids:** $\Omega_1 \supset \Omega_2 \supset \ldots \Omega_R$
- **Restriction / interpolation operators:**

Points: $\mathbf{X}_{r+1} = \mathbf{P}_r^{r+1}\mathbf{X}_r$ $\quad$ $\mathbf{X}_r = \mathbf{P}_{r+1}^{r}\mathbf{X}_{r+1} = \left(\mathbf{P}_r^{r+1}\right)^T \mathbf{X}_{r+1}$

Distances: $\Delta_{r+1} = \tilde{\mathbf{P}}_r^{r+1}\Delta_r \left(\tilde{\mathbf{P}}_r^{r+1}\right)^T$

Towards multigrid MDS

- Convex nonlinear optimization is equivalent to a nonlinear equation

$$\nabla s(\mathbf{X}) = 0 \quad \Leftrightarrow \quad \min_{\mathbf{X}} s(\mathbf{X})$$

- **Multigrid spirit**: solve problems of the form

$$\nabla s(\mathbf{X}) = \mathbf{T} \quad \Leftrightarrow \quad \min_{\mathbf{X}} s(\mathbf{X}) - \underbrace{\mathrm{trace}\left(\mathbf{X}^T \mathbf{T}\right)}_{\langle \mathbf{X}, \mathbf{T} \rangle}$$

  at different resolution levels.

- $\mathbf{T}$ - residual transferred from finer resolution levels

Modified stress

- **Problem:** the function $s(\mathbf{X}) - \mathrm{trace}\left(\mathbf{X}^T\mathbf{T}\right)$ is unbounded
- **Modified stress:** force the center of gravity of $\mathbf{X}$ to be zero

$$\hat{s}(\mathbf{X};\boldsymbol{\Delta}) \equiv \sum_{i<j}\left(d_{ij}(\mathbf{X}) - \delta_{ij}\right)^2 + \lambda \sum_{j=1}^{m}\left(\sum_{i=1}^{N} x_{ij}\right)^2$$

The modified stress is bounded

Multigrid MDS (FAS V-cycle)

$\texttt{Vcycle}\left(\mathbf{X}_r, \mathbf{T}_r, \Delta_r, K_r, K'_r\right)$

- **IF** $r = R$ (coarsest level), $\min_{\mathbf{X}_R} s_R\left(\mathbf{X}_R, \mathbf{T}_R\right)$ and return
- **ELSE**
  - Apply K_r SMACOF iterations to $s_r\left(\mathbf{X}_r, \mathbf{T}_r\right)$, return $\mathbf{X}'_r$
  - Compute $\mathbf{G}'_r = \nabla s_r\left(\mathbf{X}'_r\right)$;

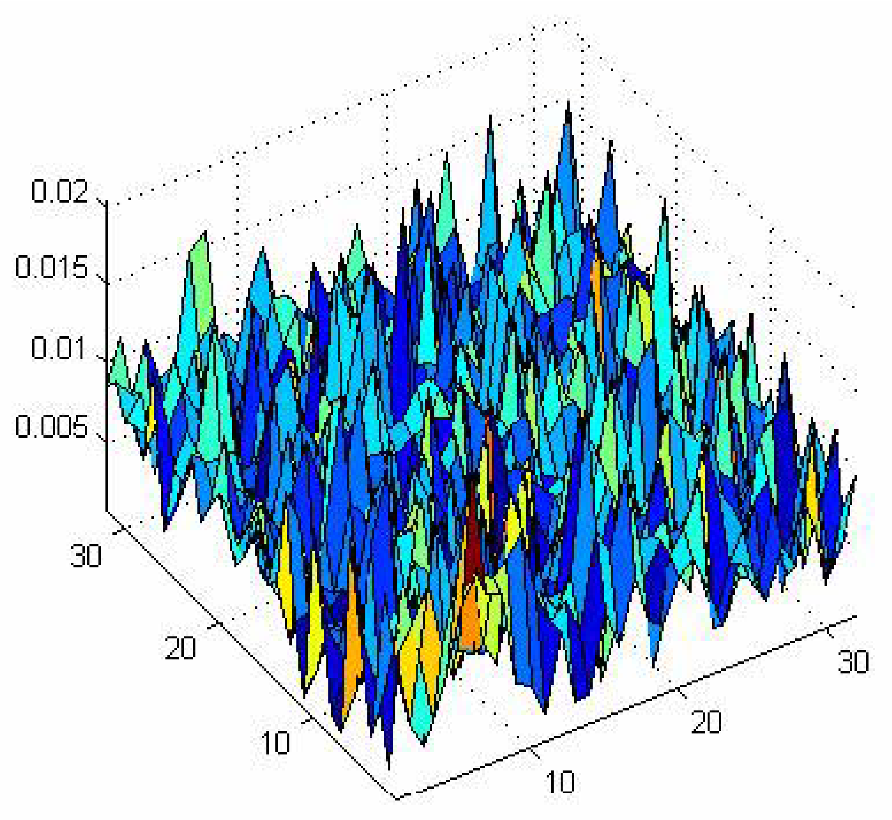
    $\mathbf{X}'_{r+1} = \mathbf{P}_r^{r+1}\mathbf{X}'_r$;

    $\mathbf{G}'_{r+1} = \nabla s_{r+1}\left(\mathbf{X}'_{r+1}\right)$;

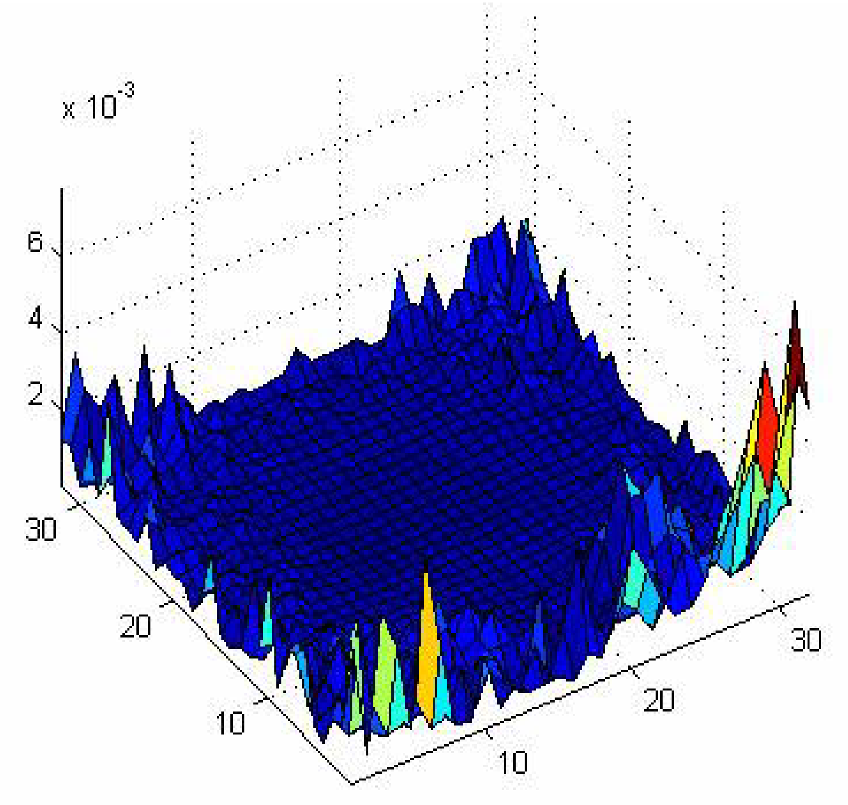
    $\mathbf{T}_{r+1} = \mathbf{G}_{r+1} - \mathbf{P}_r^{r+1}\mathbf{G}_r$
  - Apply MG on a coarser resolution:

    $\mathbf{X}''_{r+1} \leftarrow \texttt{Vcycle}\left(\mathbf{X}'_{r+1}, \mathbf{T}_{r+1}, \Delta_{r+1}, K_{r+1}, K'_{r+1}\right)$
  - Correction: $\mathbf{X}''_r \leftarrow \mathbf{X}'_r + \mathbf{P}_{r+1}^r\left(\mathbf{X}''_{r+1} - \mathbf{X}'_{r+1}\right)$
  - Apply K'_r SMACOF iterations to $s_r\left(\mathbf{X}''_r, \mathbf{T}_r\right)$, return $\mathbf{X}'''_r$

Error smoothing



BEFORE RELAXATION

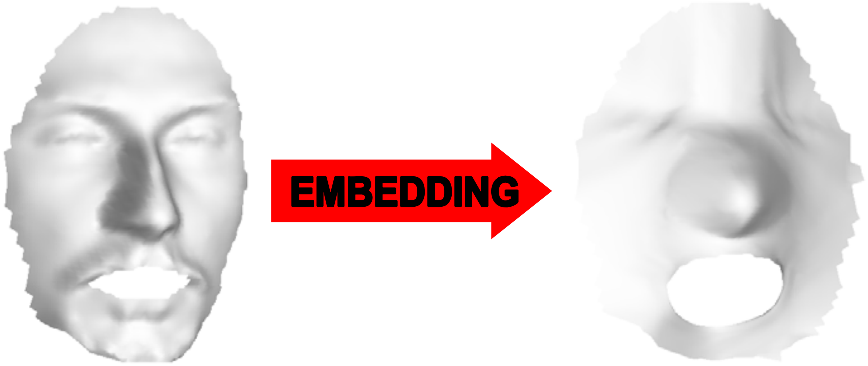


AFTER RELAXATION

Error smoothing using SMACOF relaxation

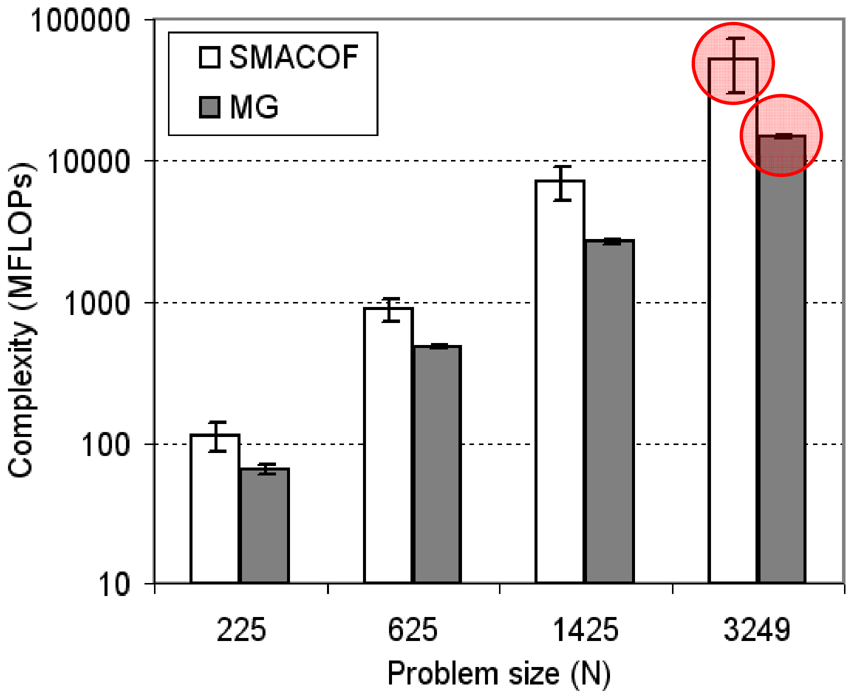
Simulations

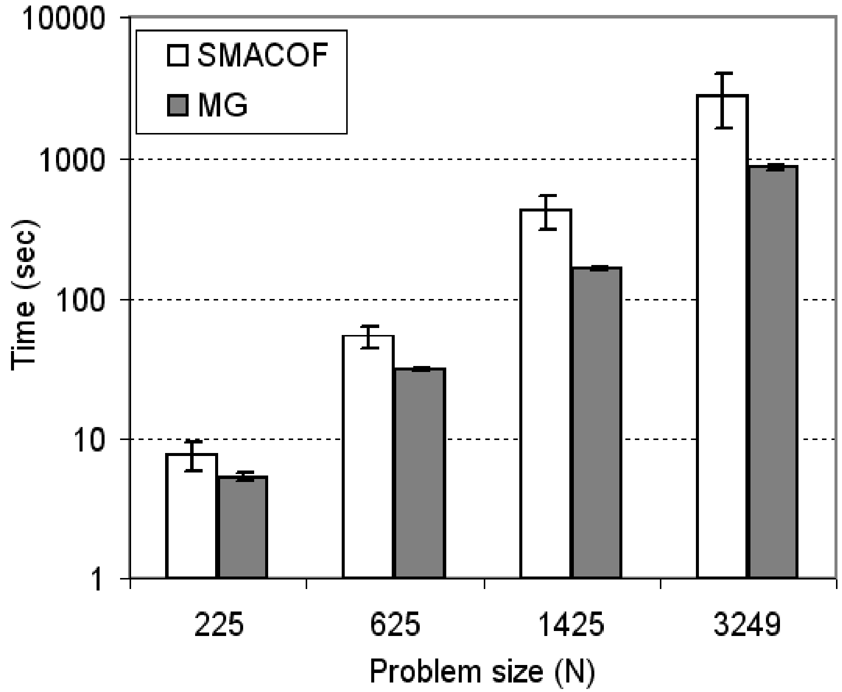
- Isometric embedding of a facial surface



- Problems of different sizes: $N = 225, 625, 1425, 3249$ points
- Different number of resolution levels: $R = 3, 4$
- Different MG cycles: **V-cycle** and **F-cycle**
- Different initialization: **original points** and **random points**

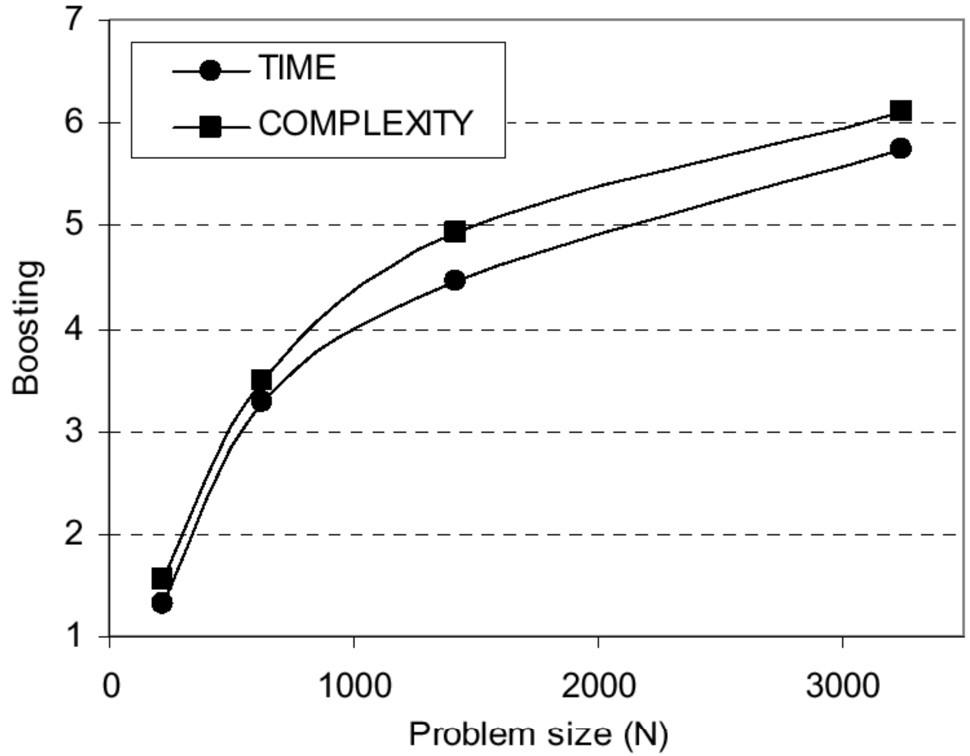
Results: random initialization





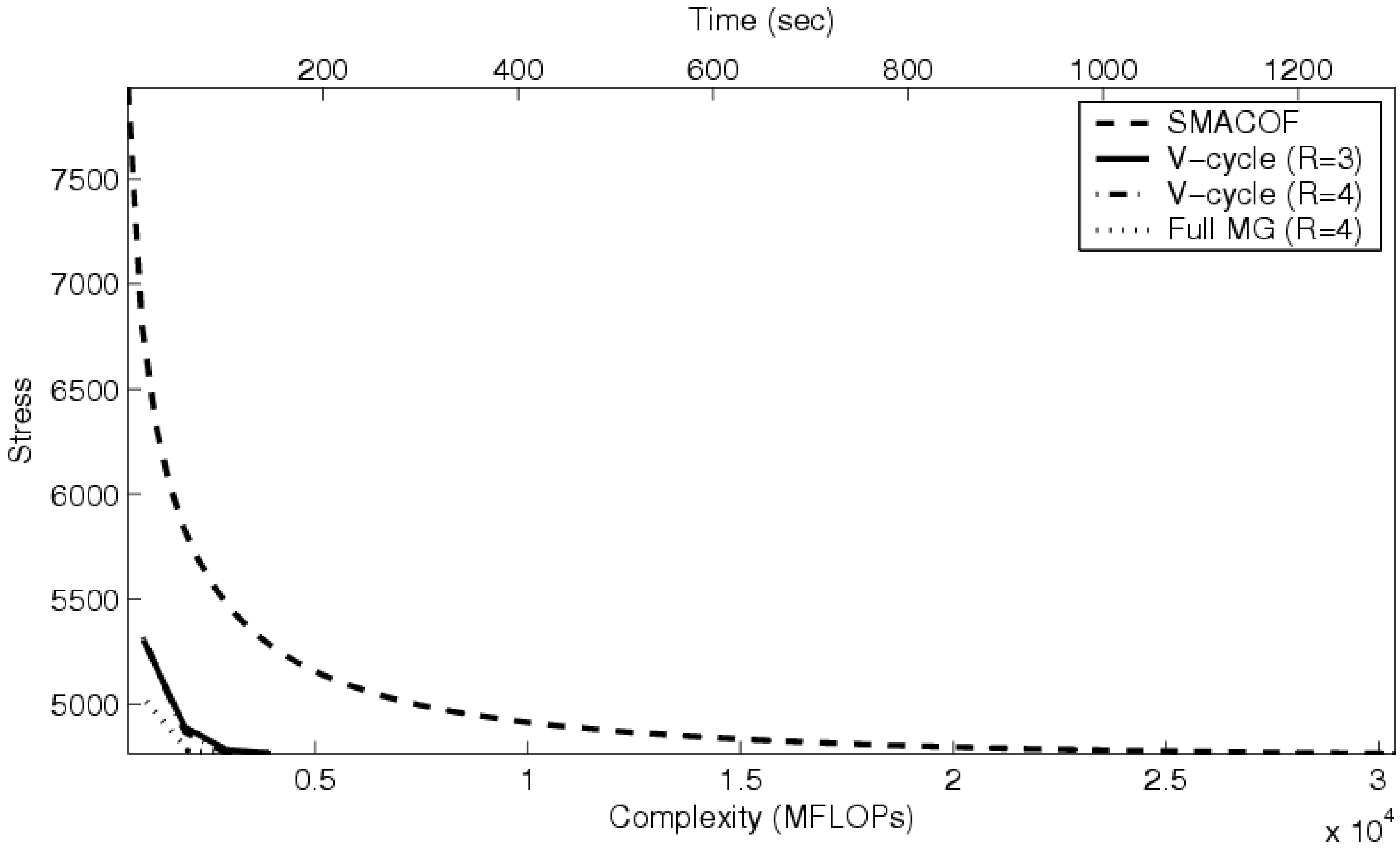
Performance of SMACOF and MG (V-cycle, 3 resolution levels) using **random initialization**

Results: Problem size



Boosting obtained by multigrid MDS (V-cycle) compared to SMACOF, 3 resolution levels, initialization by the original points

Results: different MG cycles



Convergence of different MG cycles. Initialization by the original points.

M. Bronstein, A. Bronstein, R. Kimmel, I. Yavneh, "A multigrid approach for multidimensional scaling"

Conclusions

- Multigrid MDS demonstrates **significantly better performance** compared to SMACOF (~ order of magnitude)
- The improvement is more pronounced for **large** N
- Multigrid MDS appears to be **less sensitive to initialization**

References

- A. Elad, R. Kimmel, On bending invariant signatures for surfaces, IEEE PAMI, 2003
- A. Bronstein, M. Bronstein, R. Kimmel, Expression-invariant 3D face recognition, Proc. AVBPA 2003
- A. Bronstein, M. Bronstein, R. Kimmel, Three-dimensional face recognition, IJCV, to appear
- M. Bronstein, A. Bronstein, R. Kimmel, Expression-invariant representation of faces, TR CIS-2005-01
- M. Bronstein, A. Bronstein, R. Kimmel, I. Yavneh, A multigrid approach for multidimensional scaling, MG Copper Mountain 2005
- S. Nash, A multigrid approach to discretized optimization problems, J. Optimization Methods and Software, 2002

Scale Consistent Image Completion

Michal Holtzman Gazit and Irad Yavneh

Computer Science Department
Technion - Israel Institute of Technology

[Based on: G&Y, Int. J. Multiscale Comput. Eng., 617-628, 2008]

The Problem

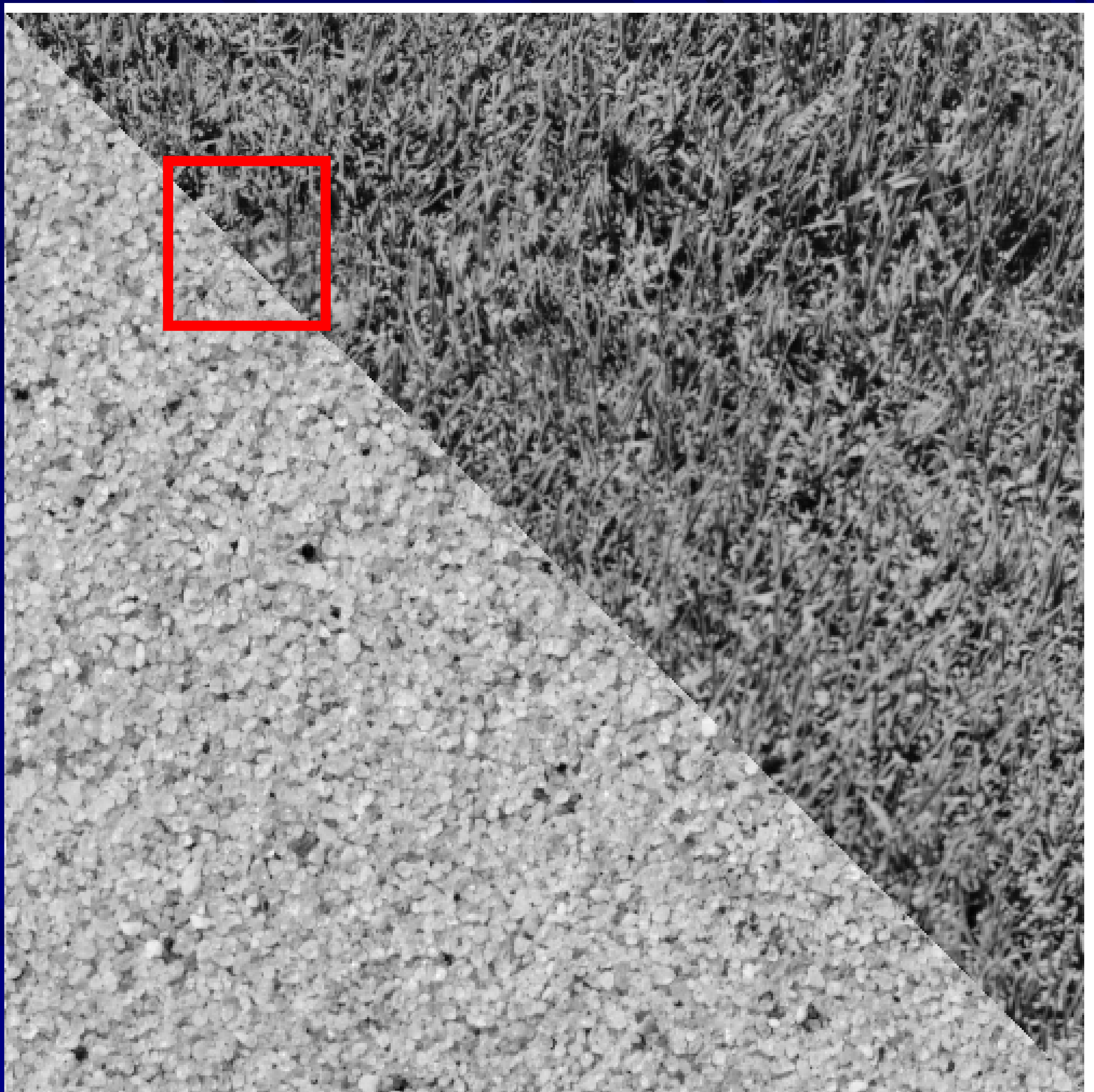
- Complete missing information in images
  - Image altered by object removal
  - Text or scratch on an image

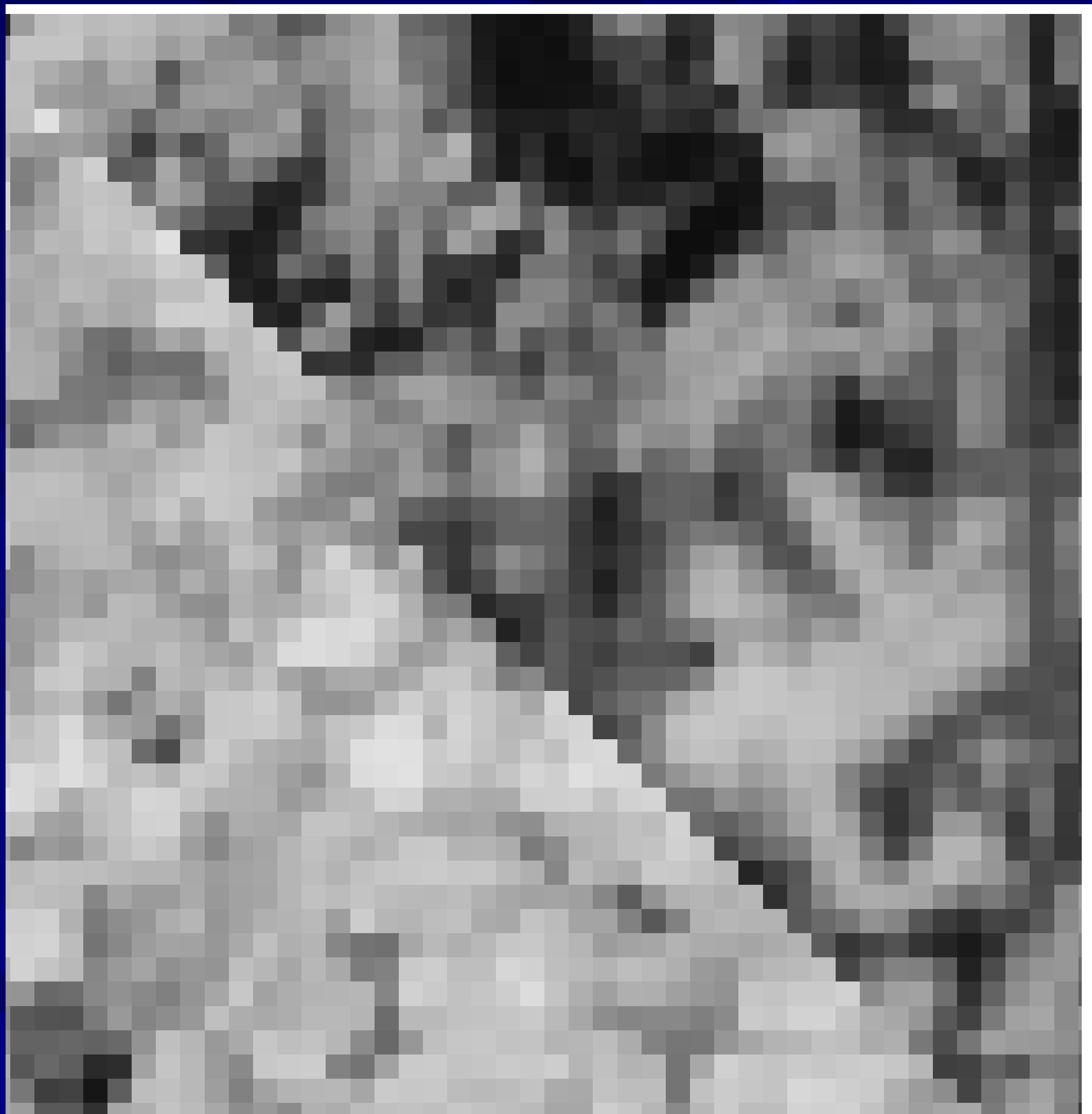
Vanishing Crane

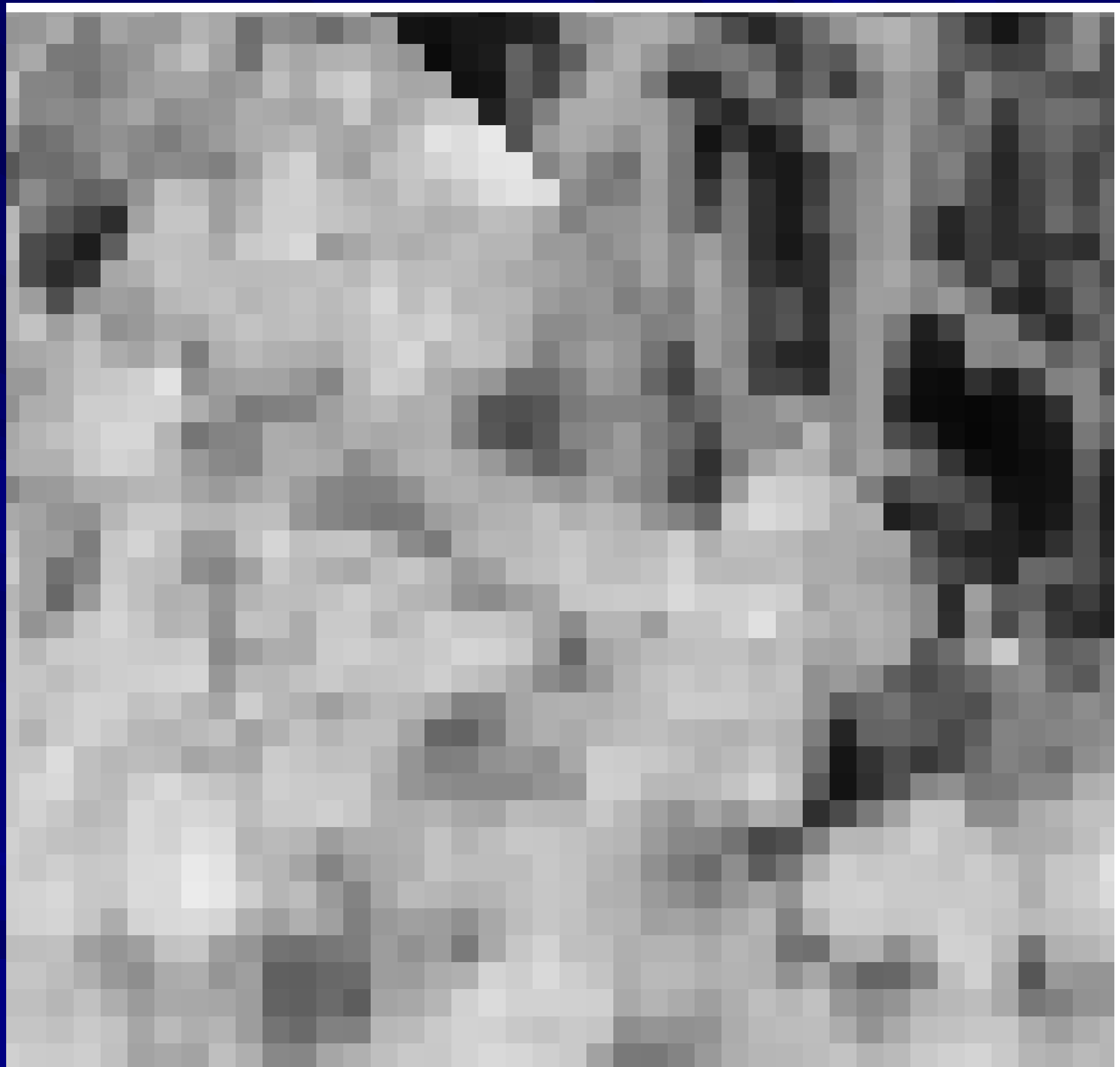
Objectives

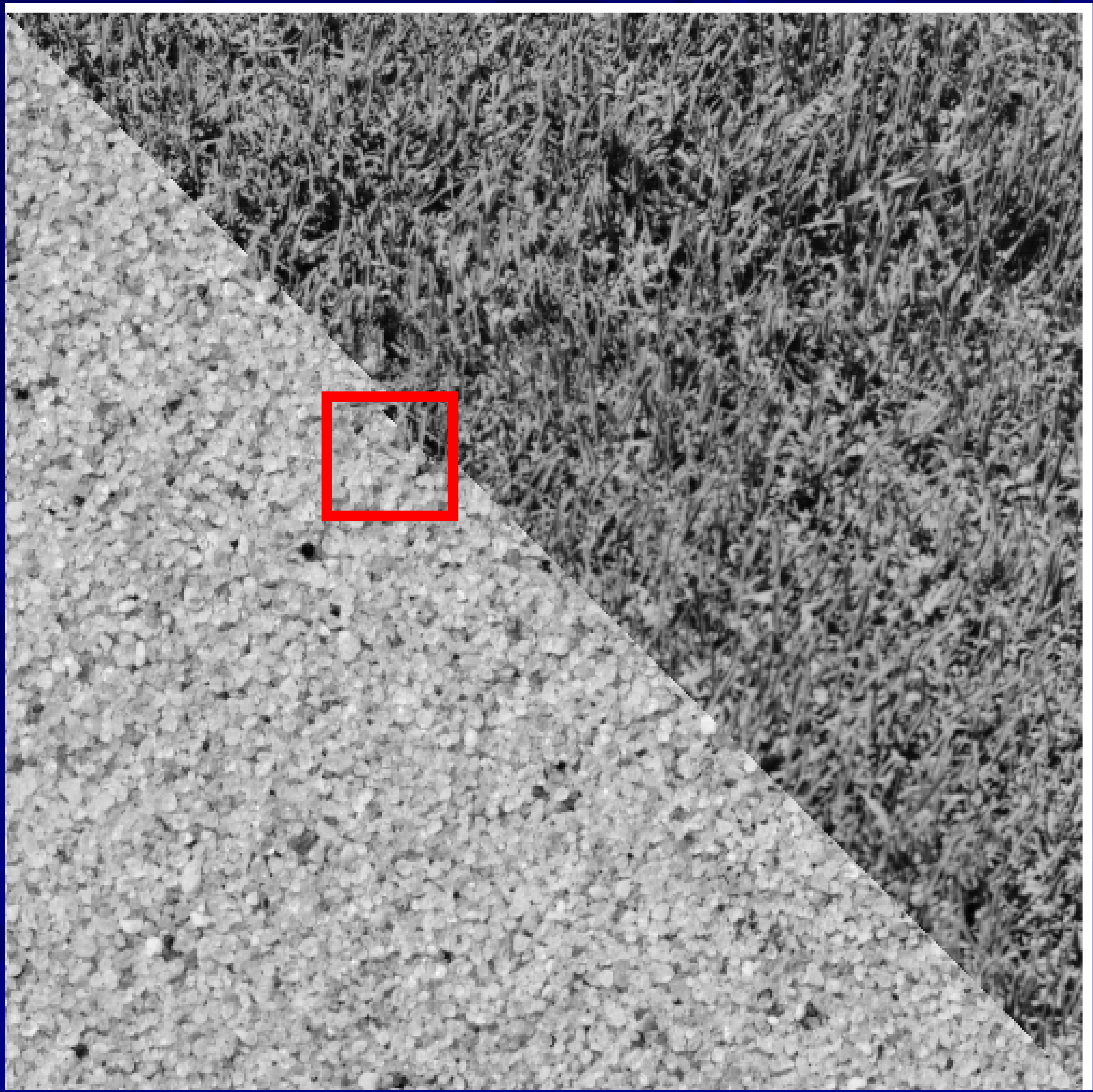
- The objective
  - To complete the image so that it will “look natural”.
- Mathematically hard to define.
  - No good objective measures of success/failure yet.
- Naturalness is multi-scaled, and ultimately requires high-level knowledge about the world.

Nevertheless, there are several good low-level approaches and many algorithms which often work well.







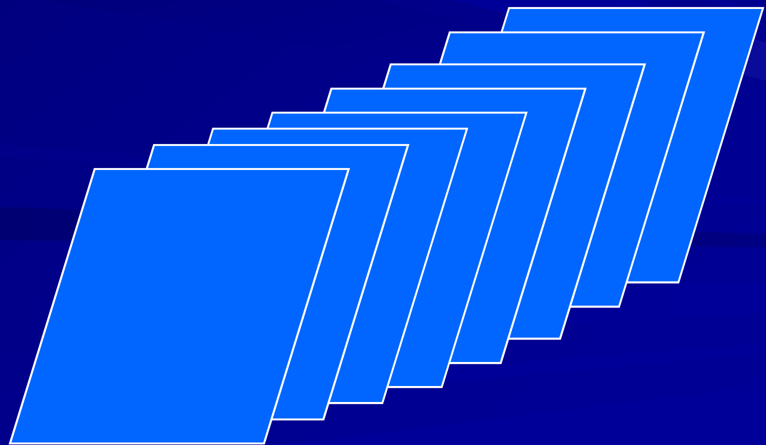


Previous work

- Inpainting Methods
  - PDE based
  - Diffusion by convolution
  - Learning image Statistics
- Texture Synthesis
  - Synthesizing one pixel at a time
  - Copying full patches onto the missing region
- Complex methods involving
  - Segmentation
  - Rotation and scaling of patch
  - Image decomposition
  - Order of filling
  - User guidance

Our Contribution

- Systematic employment of another dimension: *scale*.
- The main idea: a good completion must be *scale consistent*. That is, regardless of what our criterion of success is, the completed image must satisfy it at all scales.



Abstract Description

<u>Image</u>: $I = I(\Omega) : \Omega \to [0,1]^{d\times|\Omega|}$

<u>Domain</u> (set of pixels):

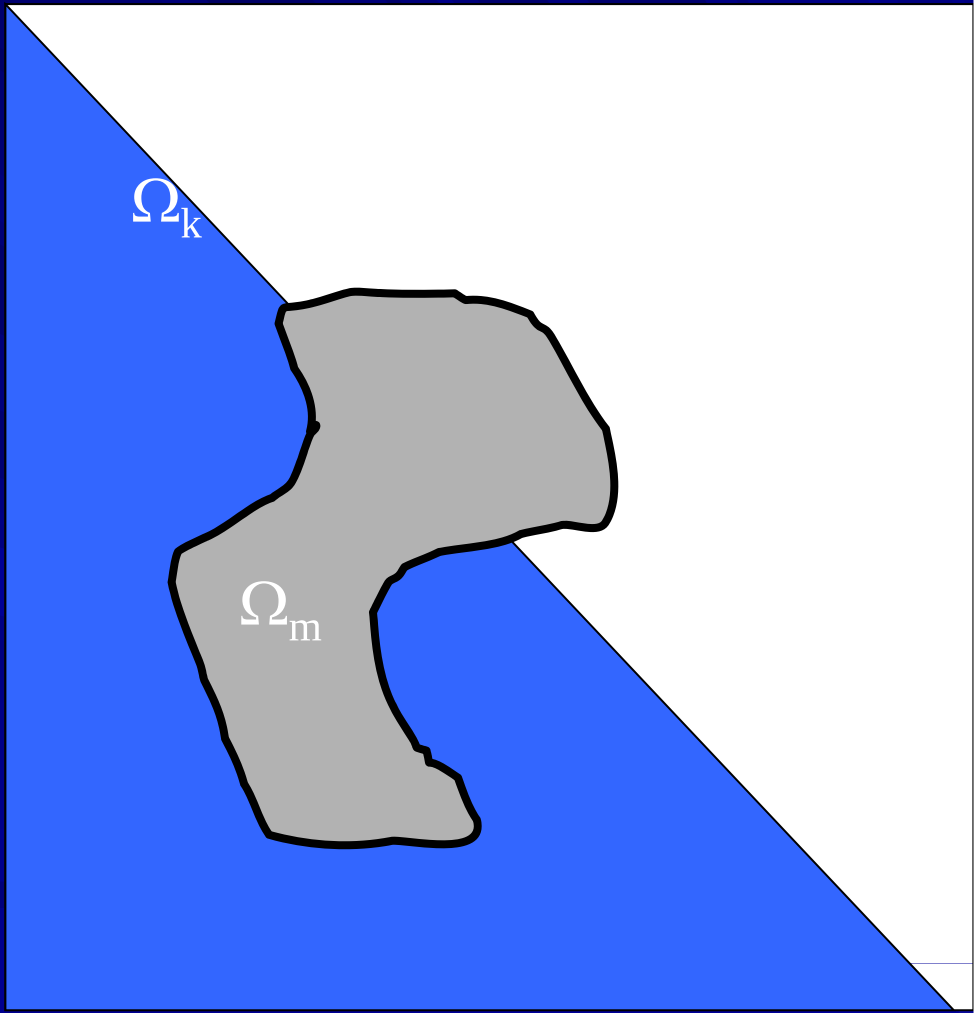
$$\Omega = \Omega_k \cup \Omega_m$$

where I is known in Ω_k but missing in Ω_m

An <u>image completion algorithm</u> is a function, $C : [0,1]^{d\times|\Omega|} \to [0,1]^{d\times|\Omega|}$

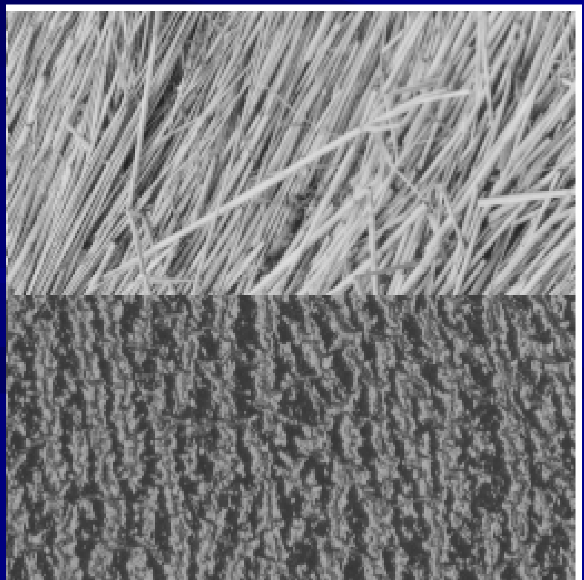
such that $\bar{I} = C(I(\Omega))$

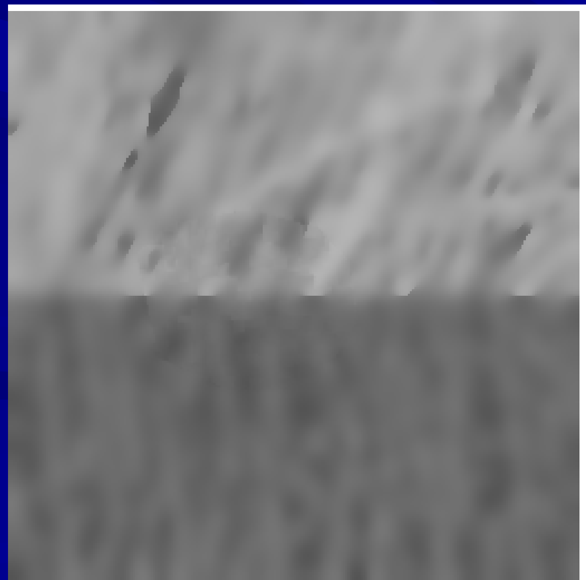
satisfies: $\bar{I}(\Omega_k) = I(\Omega_k)$

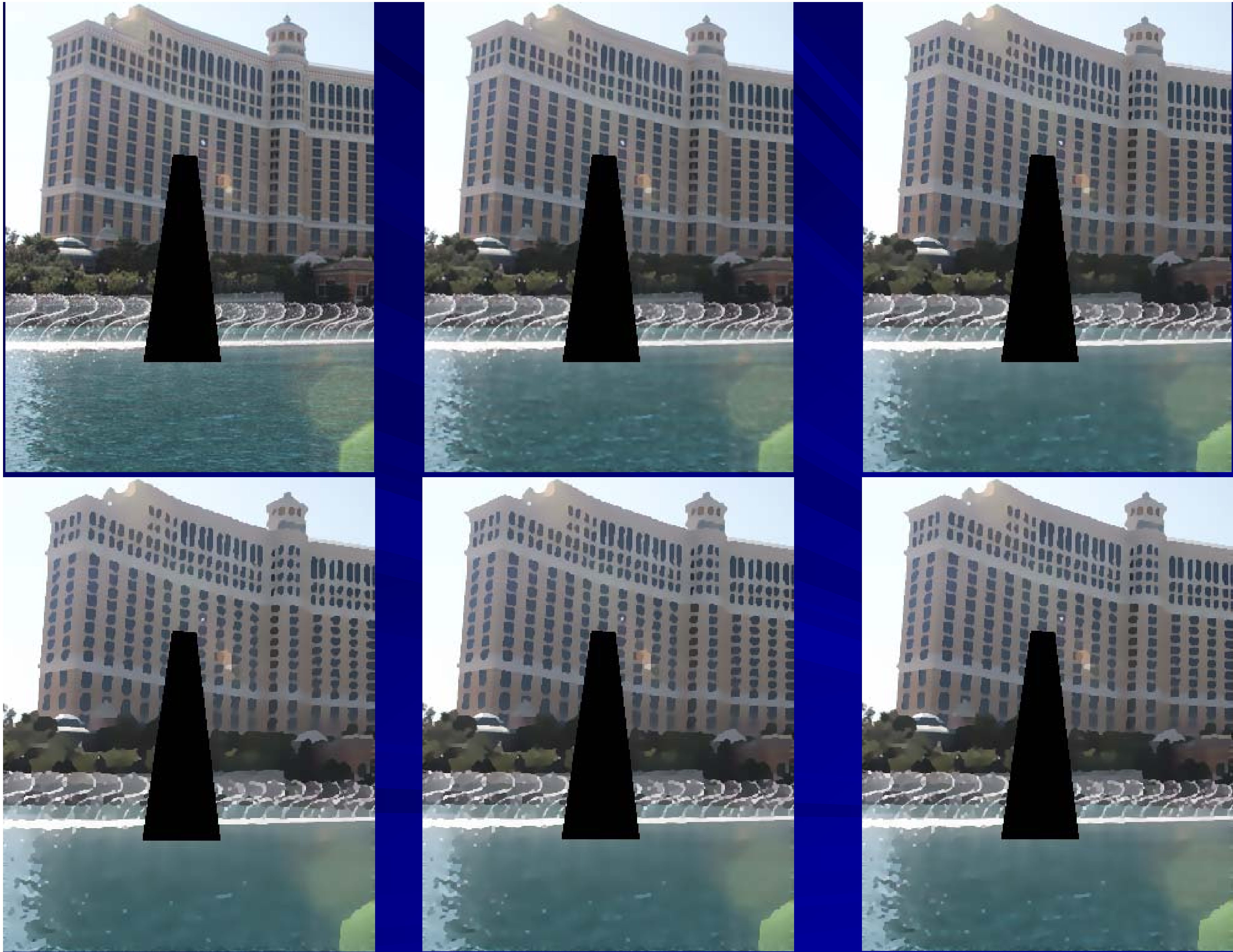


Abstract Description

A smoothing algorithm is a function,
$S:[0,1]^{d\times|\Omega|}\rightarrow[0,1]^{d\times|\Omega|}$, such that $I_S = S(I)$
is a less detailed version of I.
(T*he size of the image remains fixed*).







Scale Consistency

We say that a completion is *scale consistent* if $C\left(S\left(I\right)\right) \approx S\left(C\left(I\right)\right)$

Patch-Based Completion, C

Initialize: $\bar{I} = I$; Repeat until:

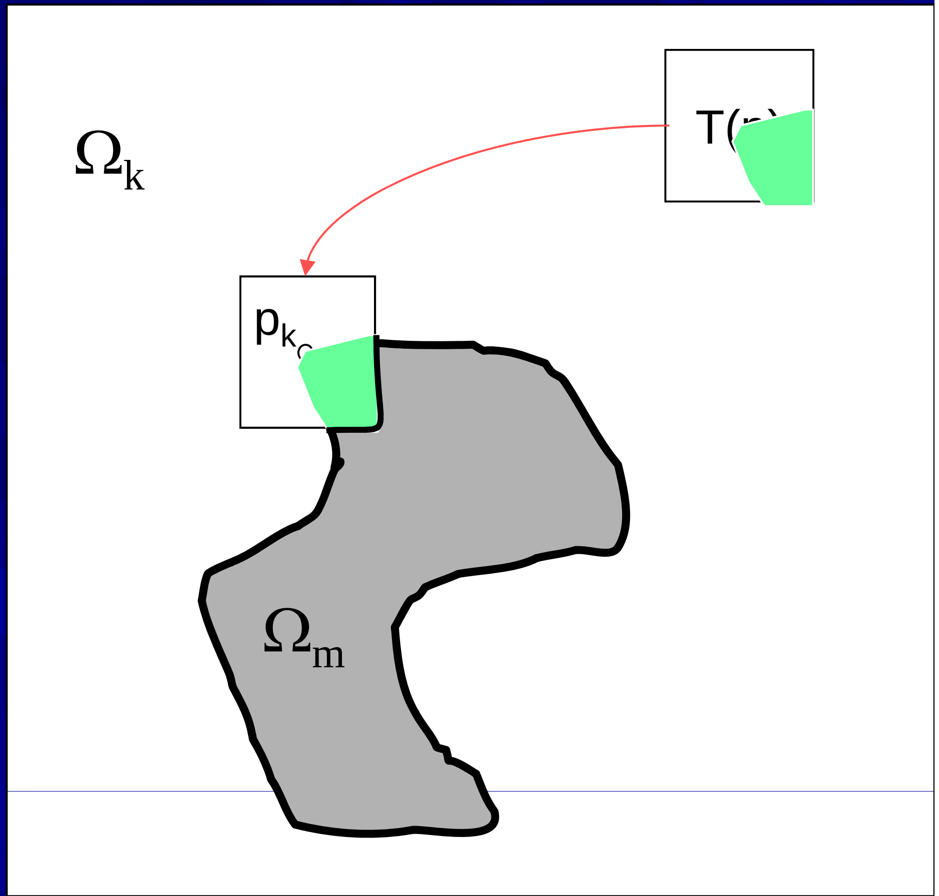
$$\Omega_m = \emptyset$$

- Choose target patch, *p*, such that

$$p_m = p \cap \Omega_m \neq \emptyset,$$

$$p_k = p \setminus p_m \neq \emptyset$$

- Choose source patch, $T(p) \subset \Omega_k$ where *T* belongs to a set of simple transformations, e.g., translations.
- Set $\bar{I}(p_m) \leftarrow \bar{I}(T(p_m))$
- Redefine $\Omega_m \leftarrow \Omega_m \setminus p_m$



Patch-Based Completion, *C*

How should the target patch, p (i.e., ordering of filling), and the source patch, $T(p)$, be chosen?

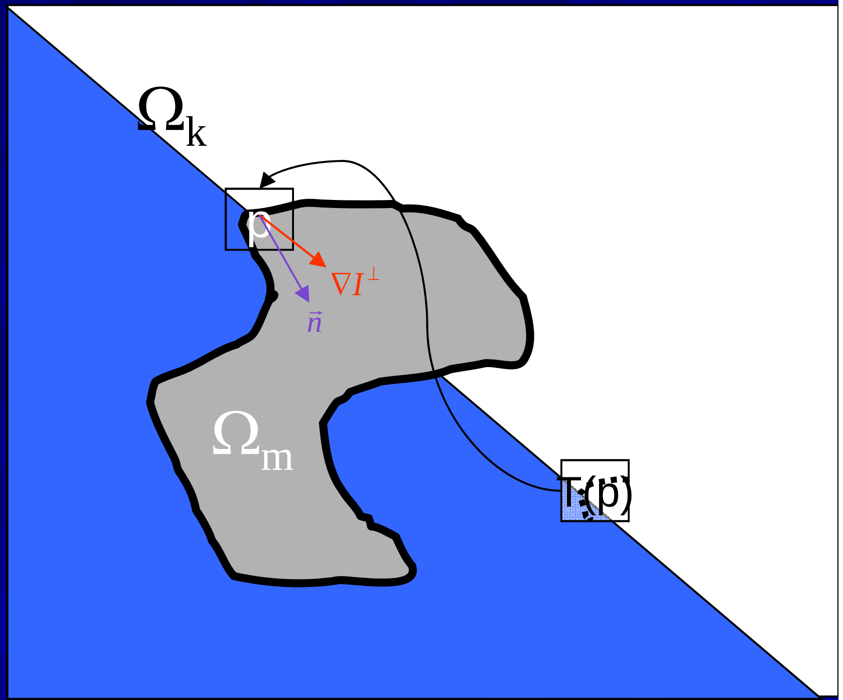
We adopt (but modify) the approach of Criminisi et al.:

A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based inpainting. *IEEE Transactions on Image Processing*, 13(9):1200–1212, 2004.

Elements of C

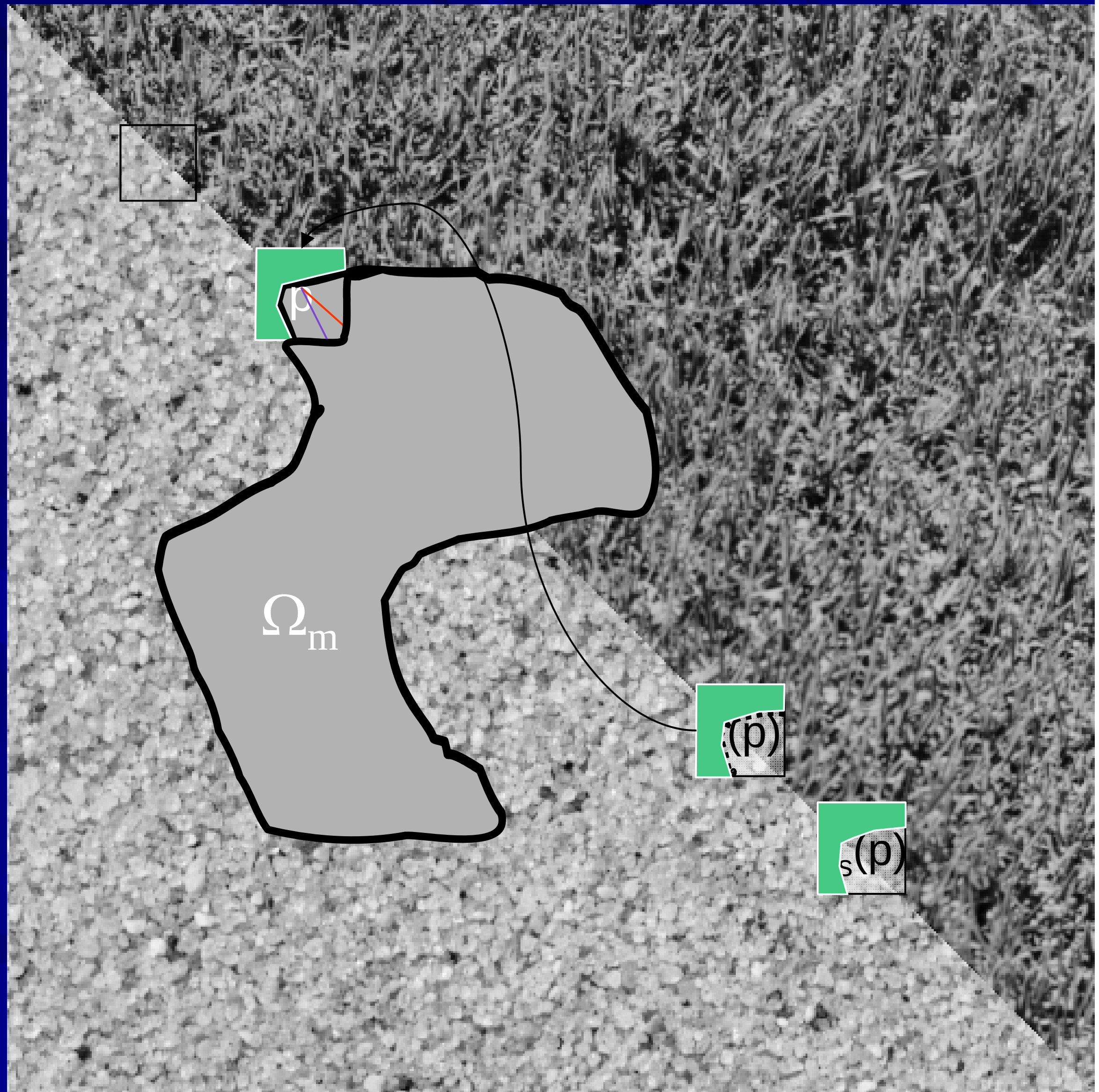
- Choosing p:
  - fix size and shape (square), and center on a boundary point of Ω_m
  - Maximize the product of
    - $|p_k|/|p|$ Confidence in that patch
    - $|\nabla I^{\perp} \cdot \vec{n}|$ The inner product between the normal to the boundary of Ω_m and the edge entering Ω_m
- Choosing $T(p)$: minimize

$$\left\| \bar{I}(p_k) - \bar{I}(T(p_k)) \right\|$$



Three Criteria $C\left(S\left(I\right)\right) \approx S\left(C\left(I\right)\right)$

1. Smoothed-image completion:
   $$\bar{I}_S\left(T_S\left(p_k\right)\right) \approx \bar{I}_S\left(p_k\right)$$
2. Detailed-image completion:
   $$\bar{I}\left(T\left(p_k\right)\right) \approx \bar{I}\left(p_k\right)$$
3. Scale consistency:
   $$\bar{I}_S\left(T\left(p\right)\right) \approx \bar{I}_S\left(p\right)$$



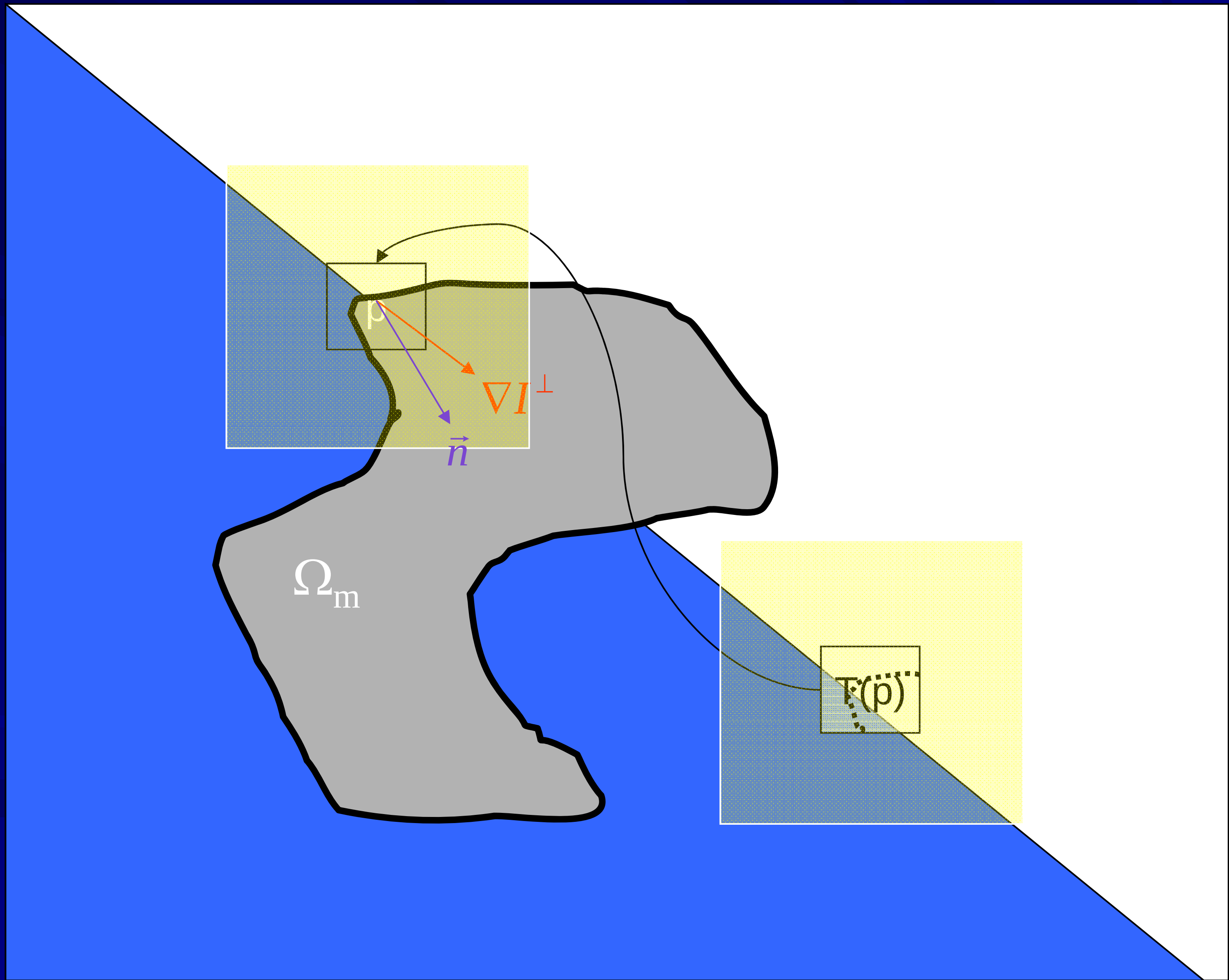
Specific Algorithm

- Generate n detail levels of I
- Complete a single patch in I_S
- Complete the same patch in I while trying to satisfy $\bar{I}\left(T\left(p_k\right)\right) \approx \bar{I}\left(p_k\right)$ and $\bar{I}_S\left(T\left(p\right)\right) \approx \bar{I}_S\left(p\right)$ simultaneously, equally weighted.
- Multi-scale: recursive, coarse-to-fine.
- Fine to Coarse:
  - The best match in the finest image is eventually used to fill the location in all the levels.

Computational Complexity

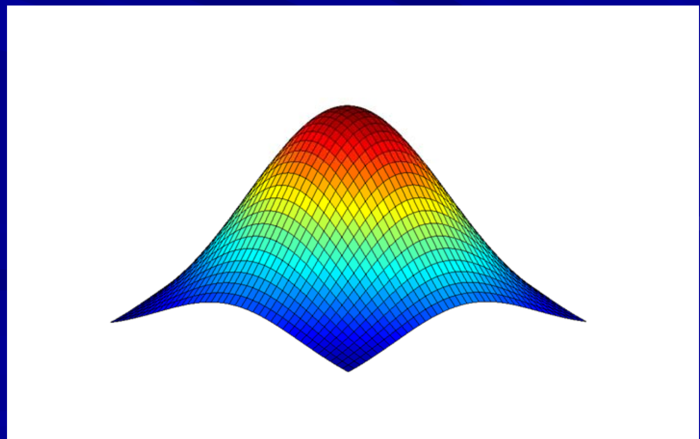
- The total complexity for n levels is only $(1+0.07(n-1))$*(Criminisi)
- Exhaustive search performed in coarse level
- Only K (~3%) best matches from coarse level are used for the finer levels for each target patch.
- Filling order is set by the coarsest level
- Each level costs 7% of the computational complexity of the coarsest level.

Single-Scale Developments



Single-Scale Developments

- Region consistent completion
  - In choosing the best matching patch, take into account the region surrounding *p*.
  - Among the *N* best matching patches choose one which has a similar surrounding to the surrounding of *p*.
  - Give decreasing weight to the pixels far from the center point (due to lower relevance).

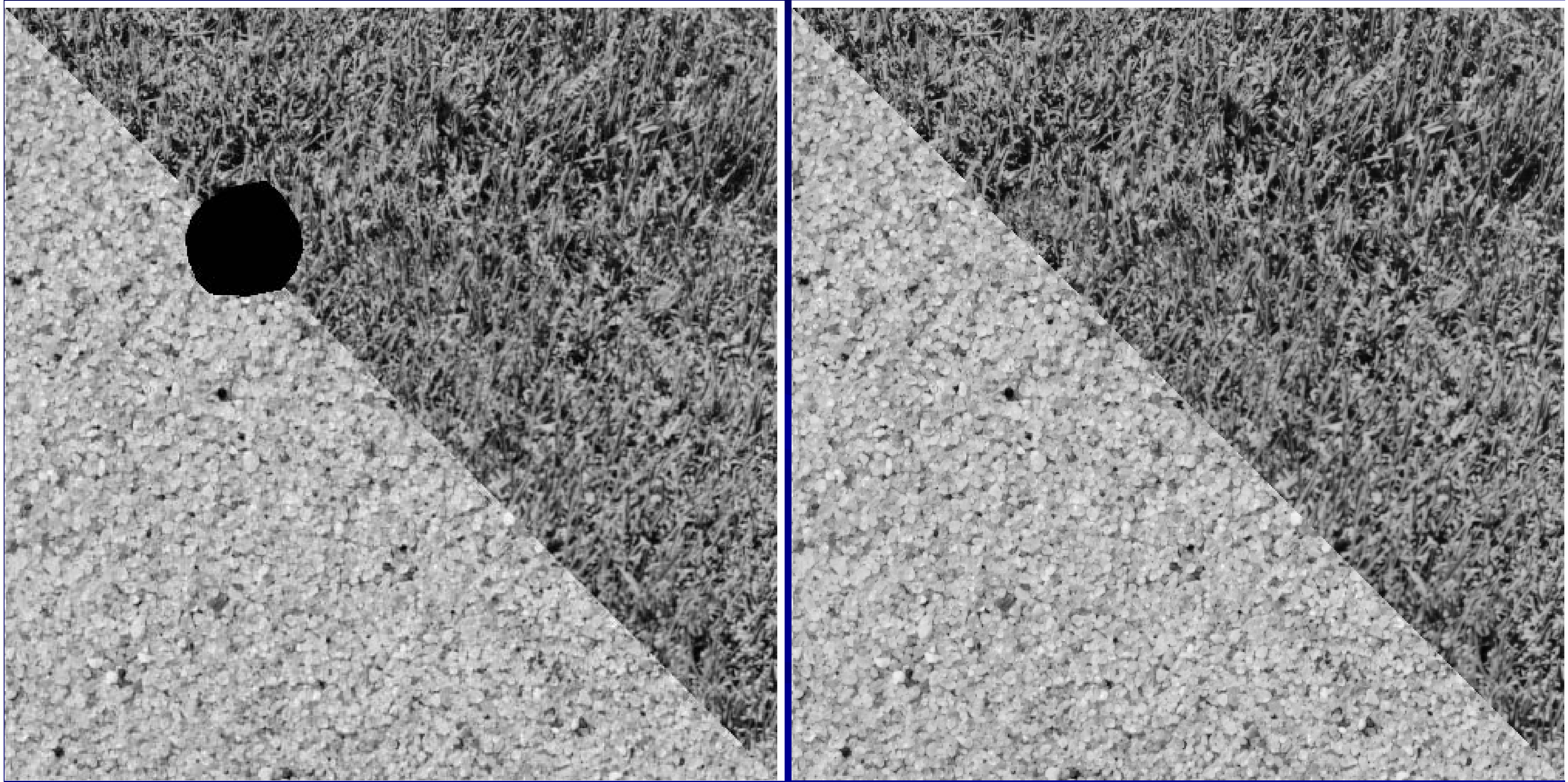


Experiments

- Systematic comparison on a synthetic image of 500x500 pixels containing 2 textures.
- To add randomness, tested 50 locations of the missing region
- Subjective grading
  - Q=1 visible defect
  - Q=2 good (slight defects)
  - Q=3 excellent
- Compared SCIC to Criminisi.

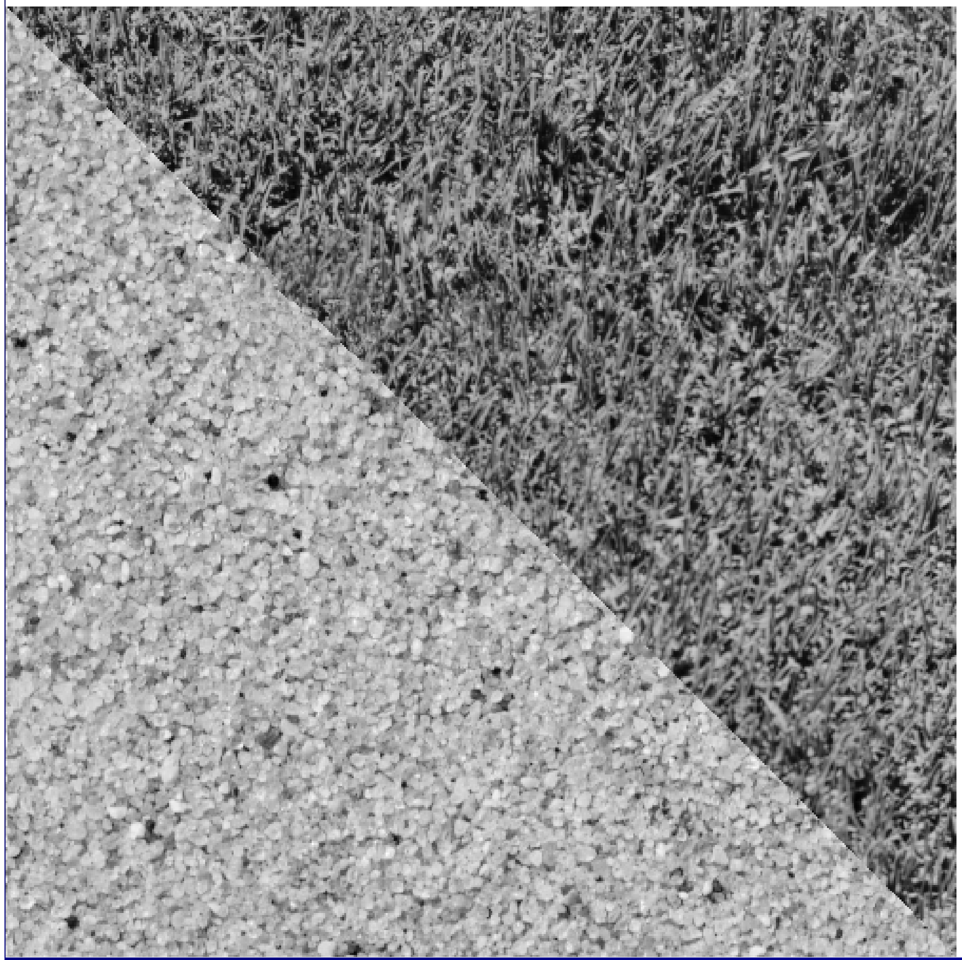
Examples: Quality

3

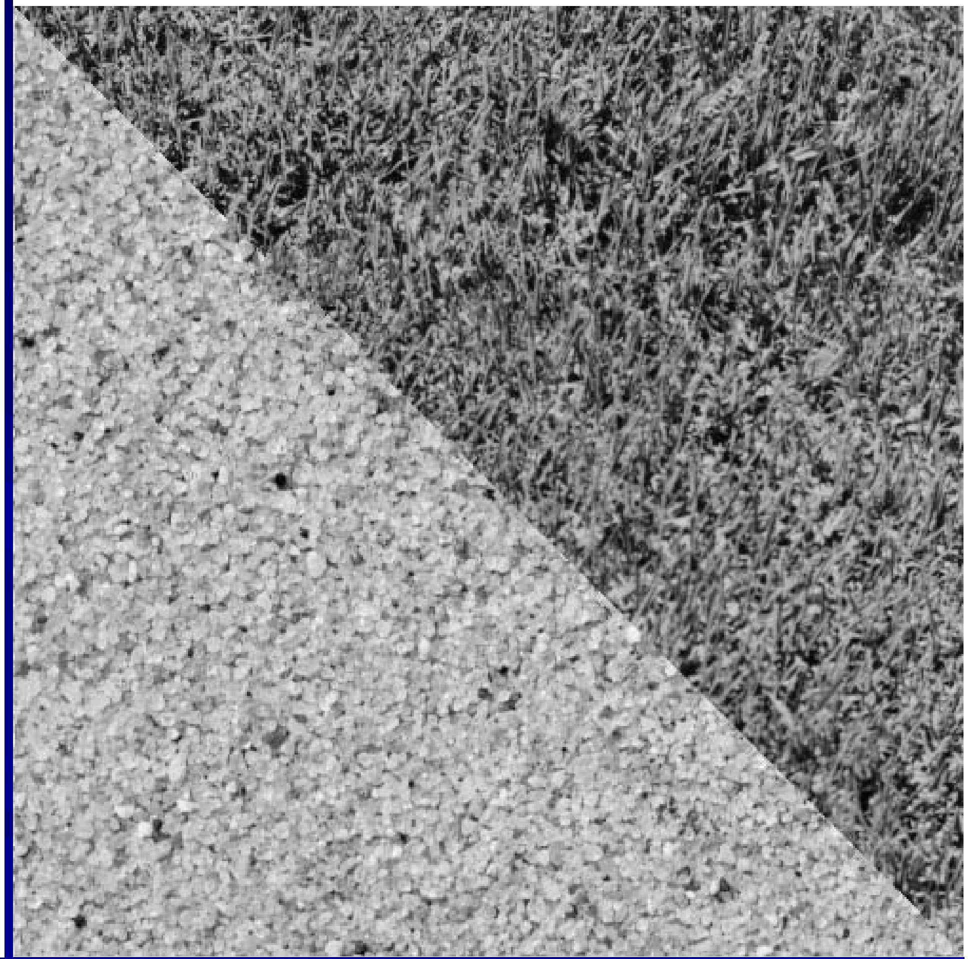


Examples: Quality

2



1



Examples: Comparison

Q	Criminisi	SCIC
1	56%	18%
2	36%	18%
3	8%	64%
Mean Score	1.52	2.46

Examples: Input Image

Criminisi et al.

SCIC

Original

Criminisi et al.

SCIC

Original

Criminisi et al.

SCIC

Criminisi et al.

SCIC

Conclusions

- Scale consistency boosts the performance of an existing patch-based completion algorithm substantially
- Fine to coarse and coarse to fine information flow
- Region Consistency
- Computational complexity – a fraction more than single scale
- Future research: Other image applications that use scale consistency.