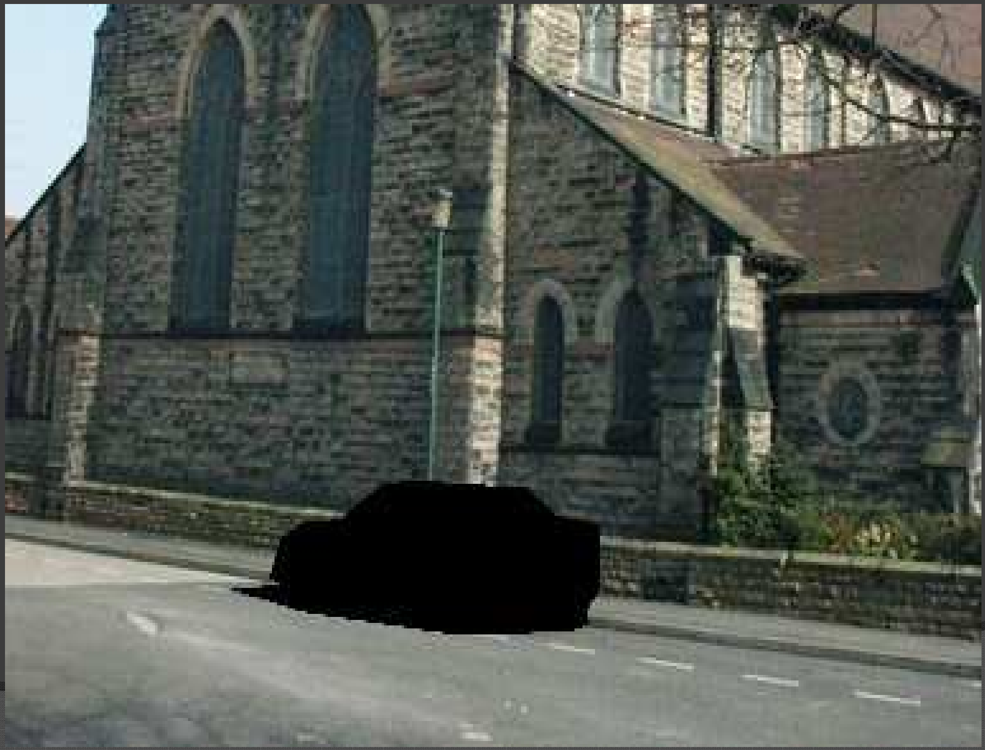
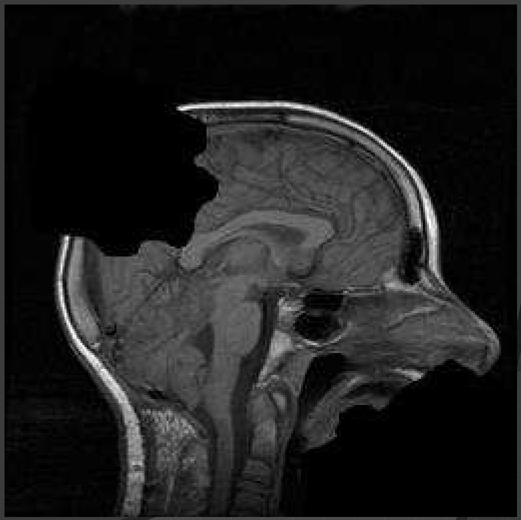
Scale Consistent Image Completion

Michal Holtzman Gazit
Irad Yavneh

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

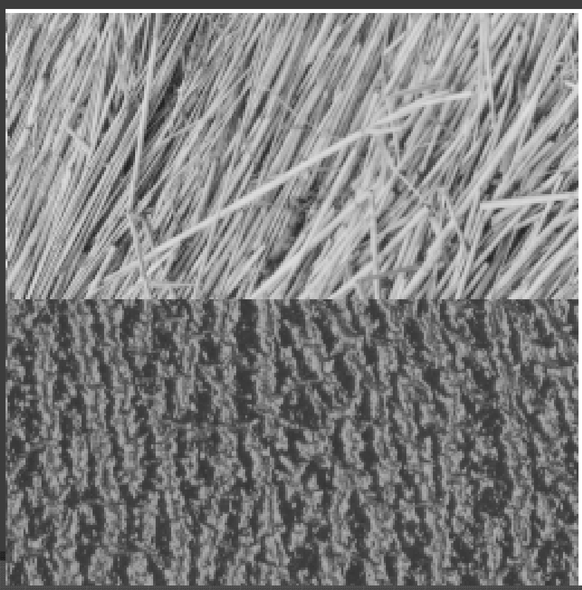
Our Contribution

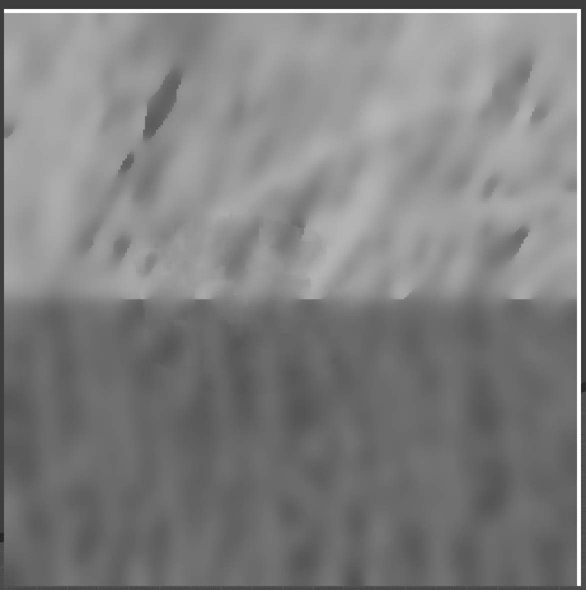
- Systematic employment of another dimension: *scale*.
- The main idea:
  - A "good" completion must be *scale consistent*.
  - Criterion of success - must be satisfied it at all scales.

Smoothing

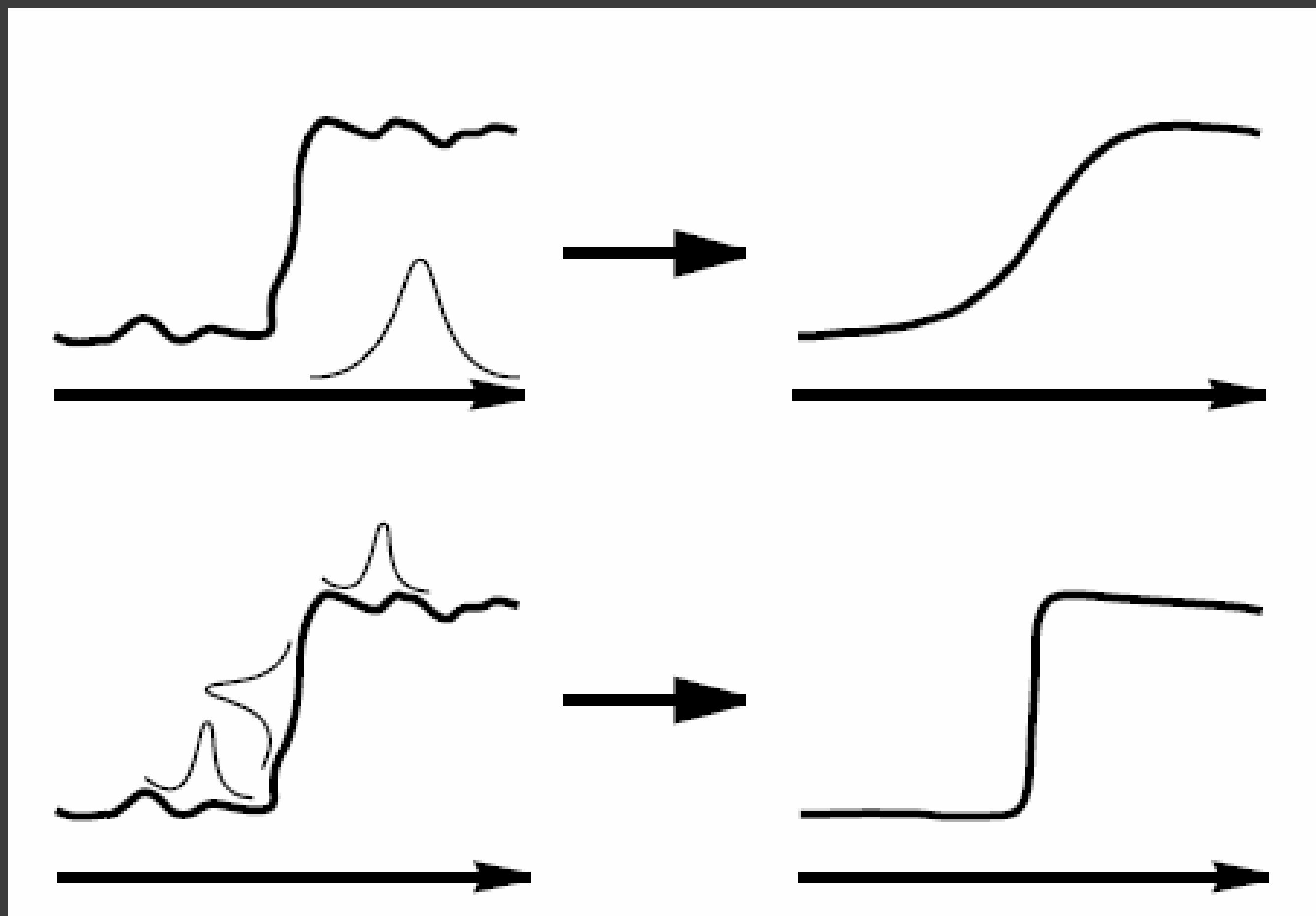
A <u>smoothing algorithm</u> is a function, $S:[0,1]^{d\times|\Omega|}\rightarrow[0,1]^{d\times|\Omega|}$, such that $I_S=S(I)$ is a less detailed version of I.

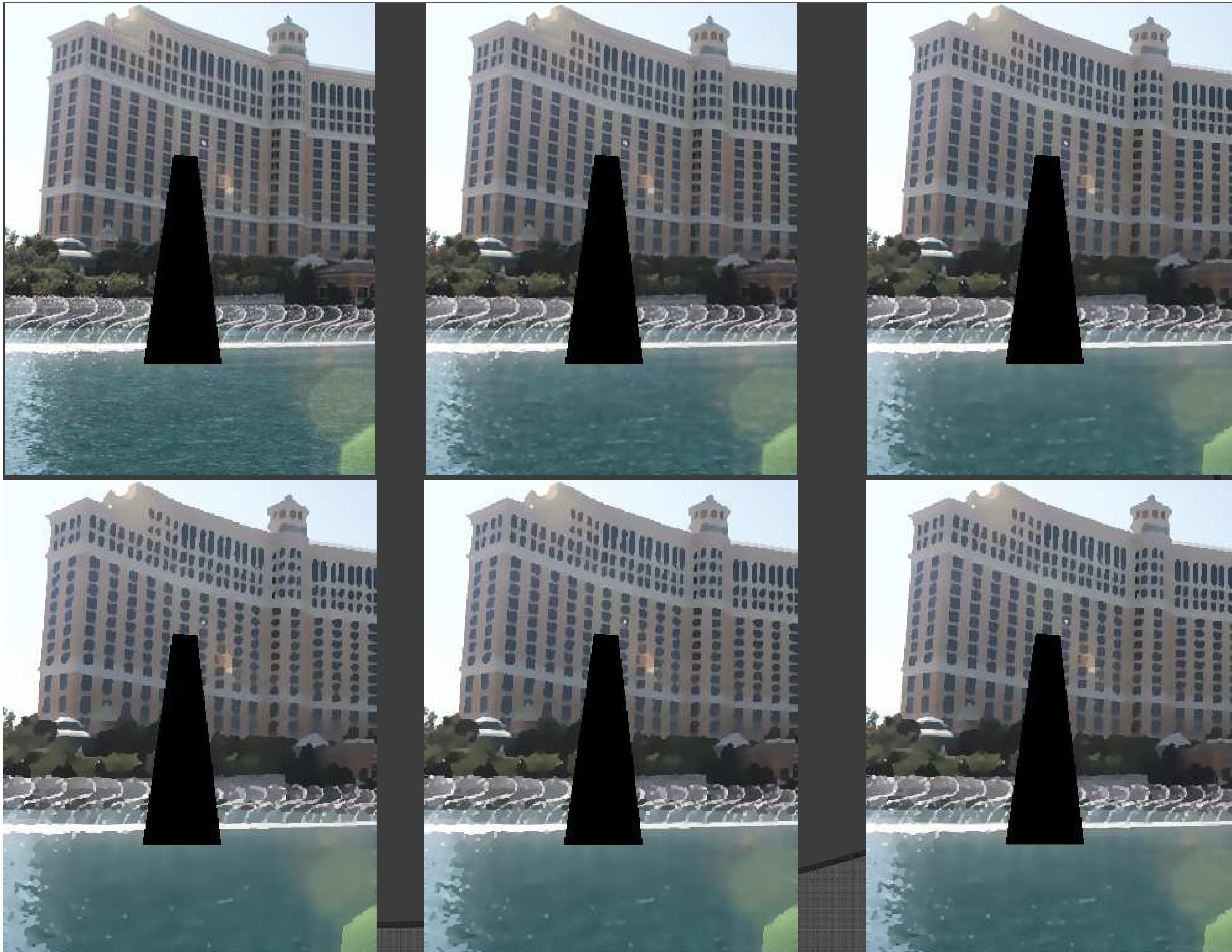
(T*he size of the image remains fixed*).





Edge Preserving Filter





Scale Consistency

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

Patch-Based Completion, C

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

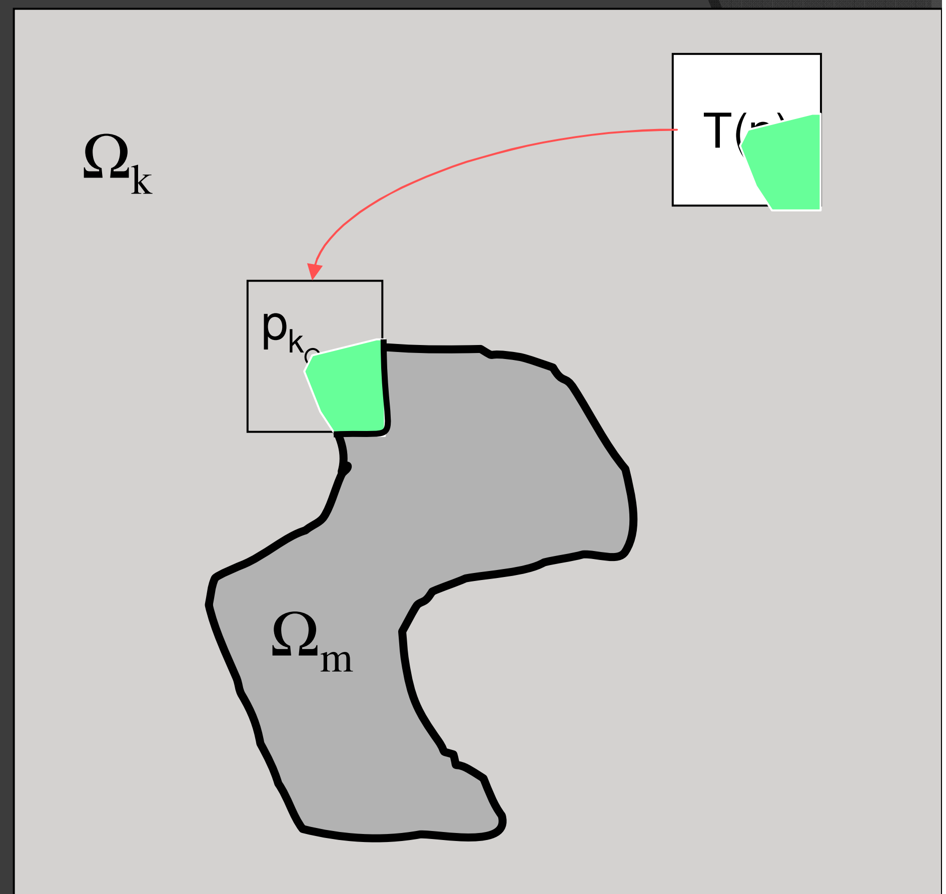
- Choose target patch, *p*, such that $p_m = p \cap \Omega_m \neq \varnothing,$

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

- 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*

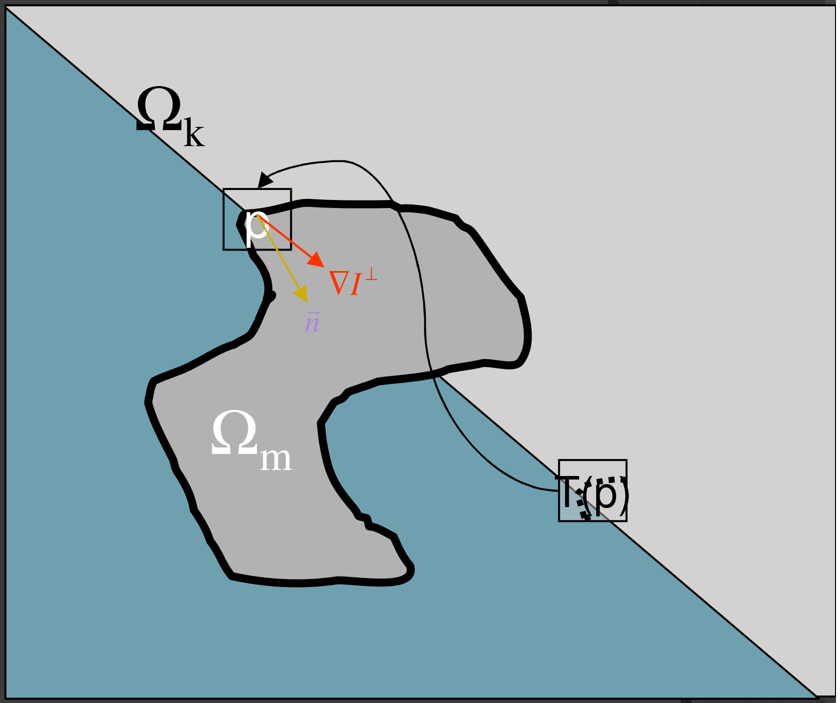
How should the target patch, *p* (i.e., ordering of filling), and the source patch, *T*(*p*), be chosen?

We adopt (but modify) the approach of Criminisi [1]

[1] 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 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

$$C(S(I)) \approx S(C(I))$$

1. Smoothed-image completion:

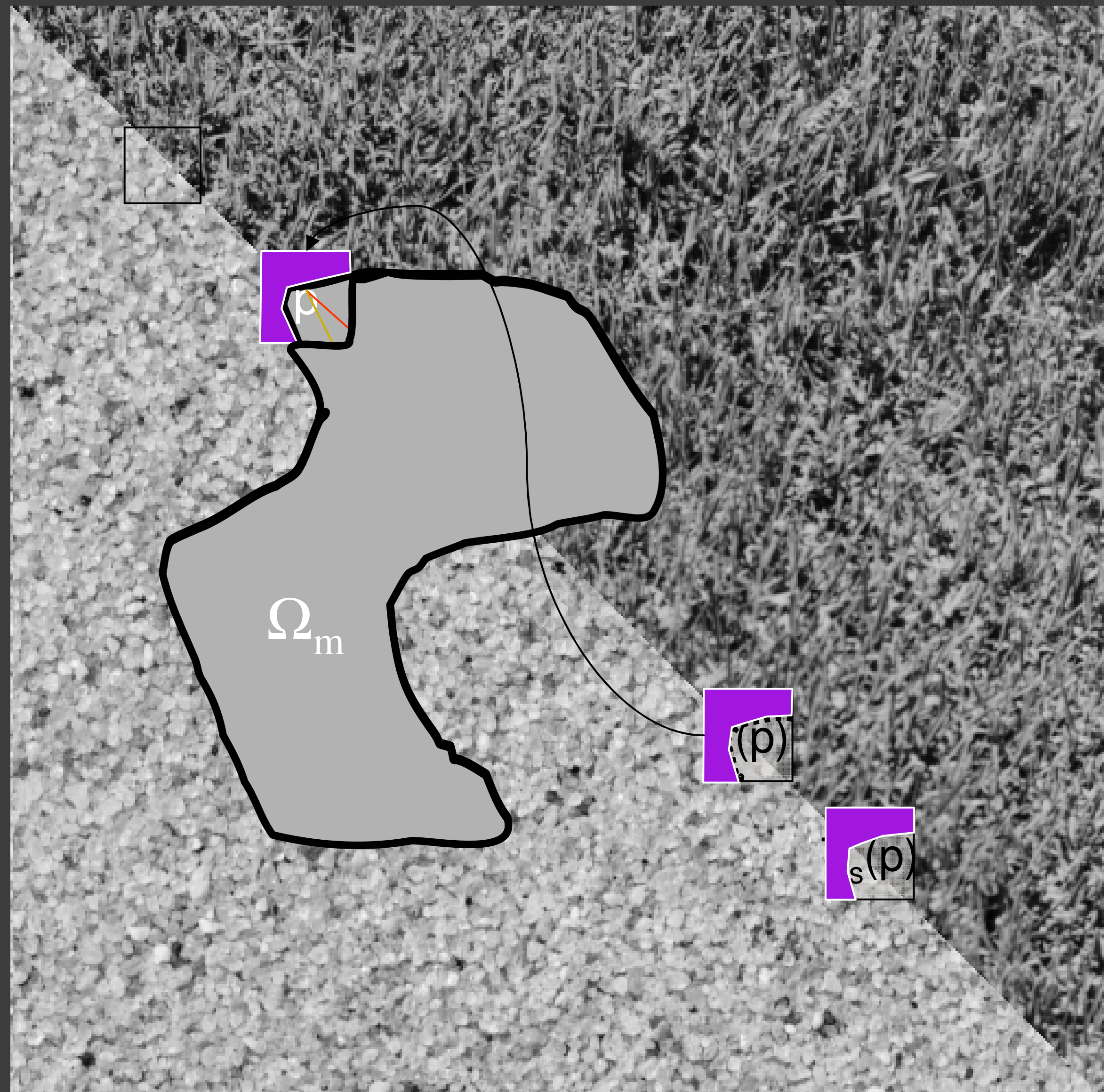
$$\bar{I}_S(T_S(p_k)) \approx \bar{I}_S(p_k)$$

2. Detailed-image completion:

$$\bar{I}(T(p_k)) \approx \bar{I}(p_k)$$

3. Scale consistency:

$$\bar{I}_S(T(p)) \approx \bar{I}_S(p)$$



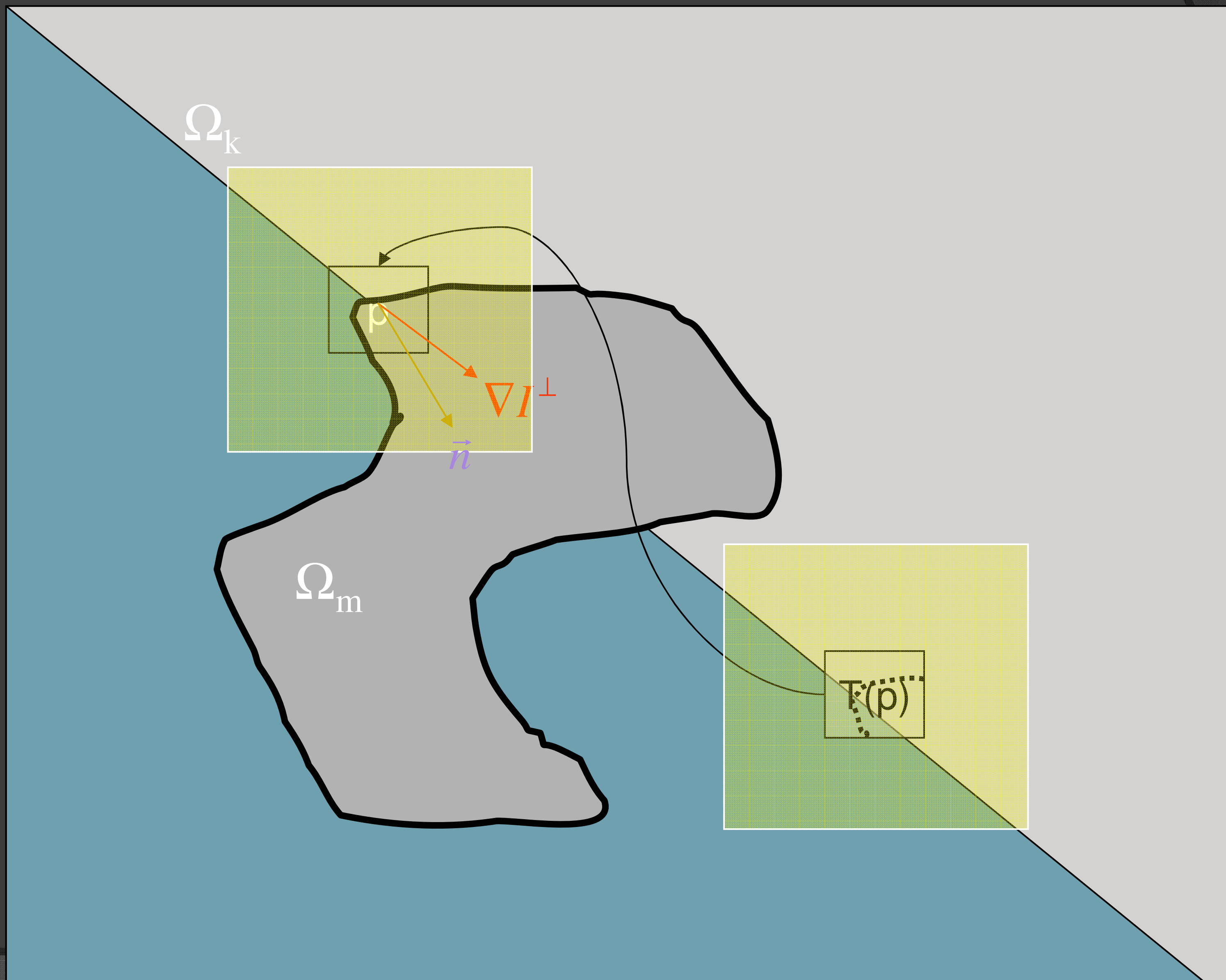
Specific Algorithm

- Generate n detail levels of I
- Complete a single patch in I_S
- Complete the same patch in I while trying to satisfy $\bar{I}(T(p_k)) \approx \bar{I}(p_k)$ and $\bar{I}_S(T(p)) \approx \bar{I}_S(p)$ 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

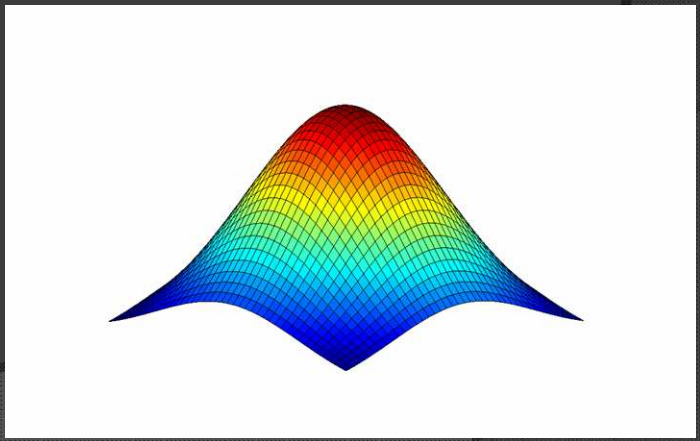
- Exhaustive search performed in coarse level
- Only K (~3%) best matches from coarse level are used for the finer levels for each target patch.
- Each level costs 7% of the computational complexity of the coarsest level
- The total complexity for n levels is only (1+0.07(n -1))*(Criminisi)
- Filling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

Single-Scale Consistency



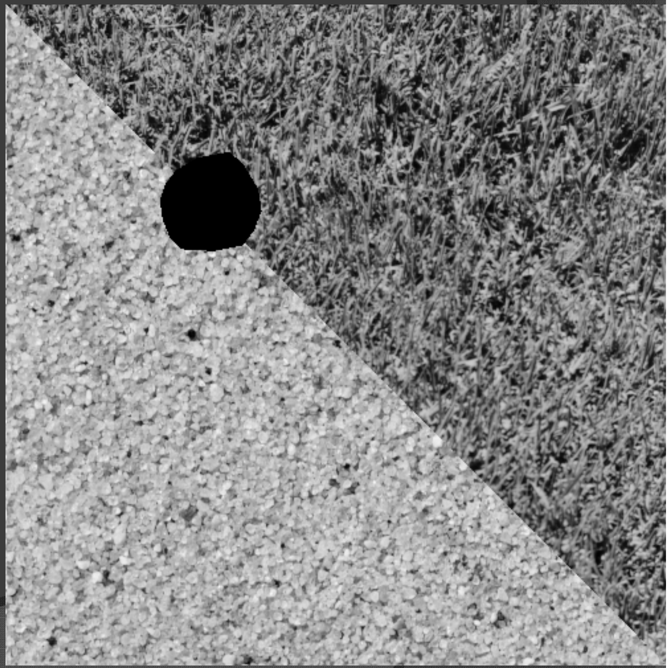
Region Consistency

- 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).



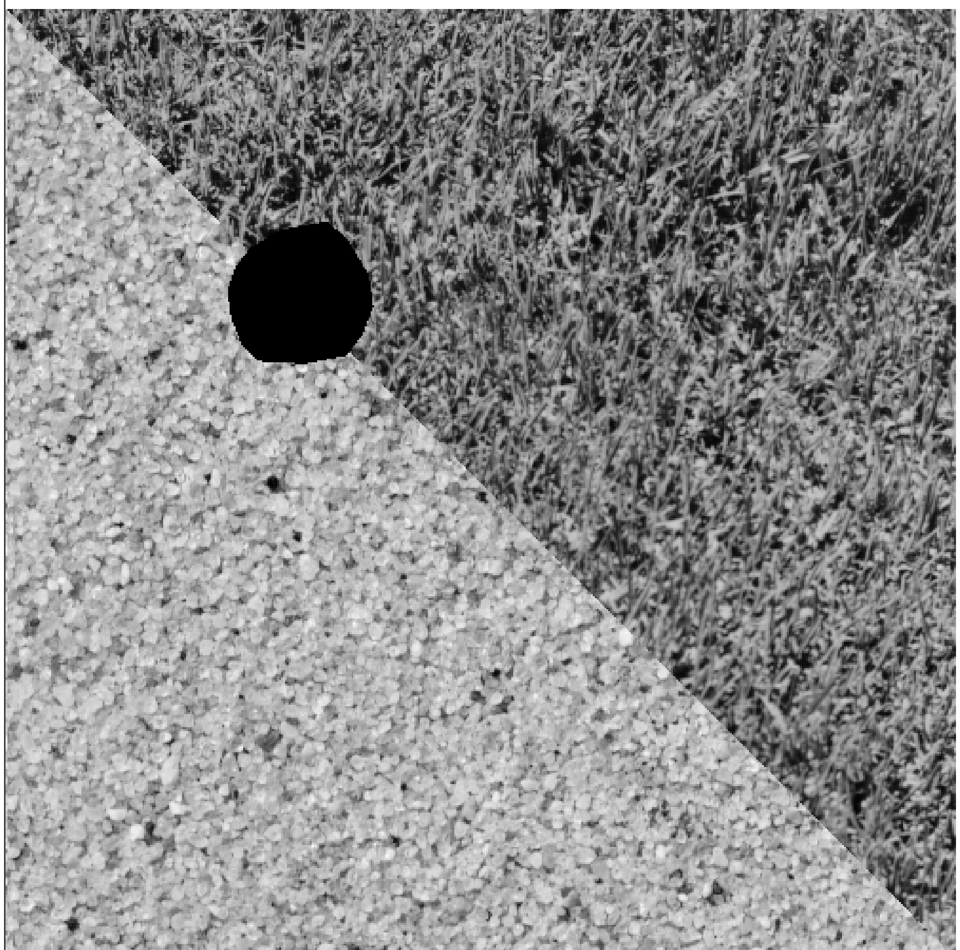
Experiments

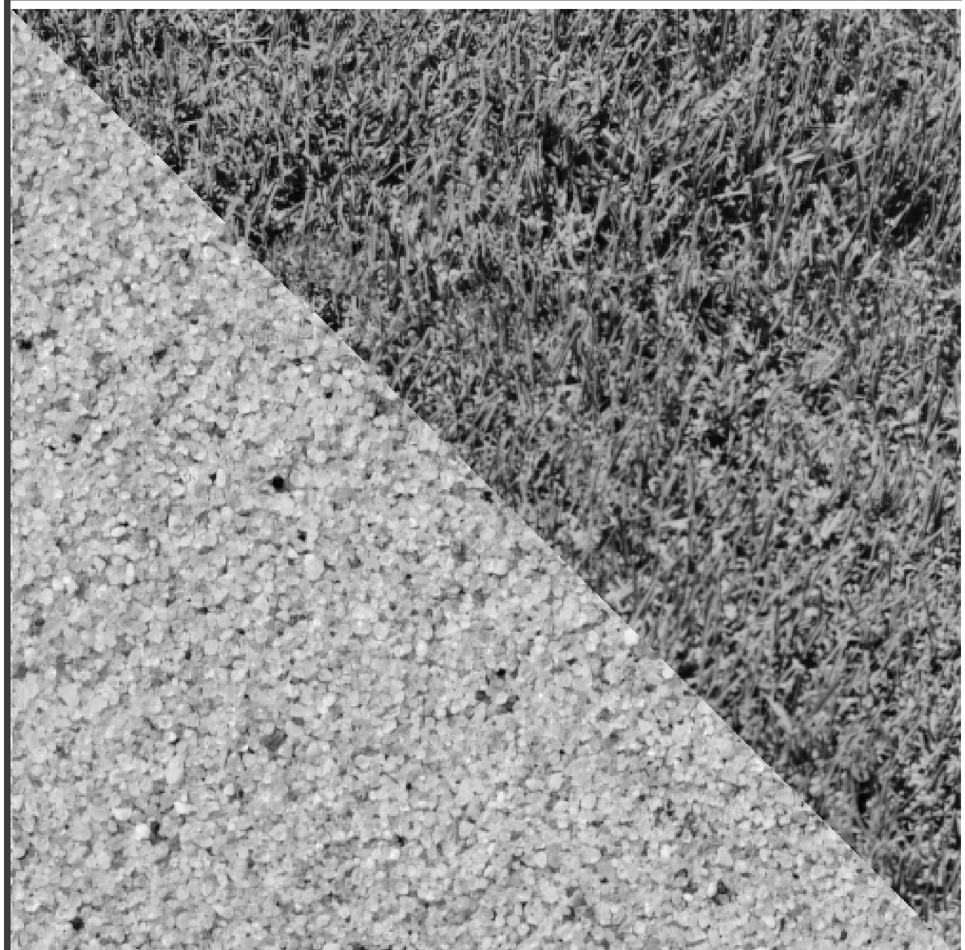
- Systematic comparison on a synthetic image of 500x500 pixels containing 2 textures.
- To add randomness, tested 50 locations of the missing region
- Subjective grading
  - Q=1 visible defect
  - Q=2 good (slight defects)
  - Q=3 excellent
- Compared SCIC to Criminisi.



Examples: Quality

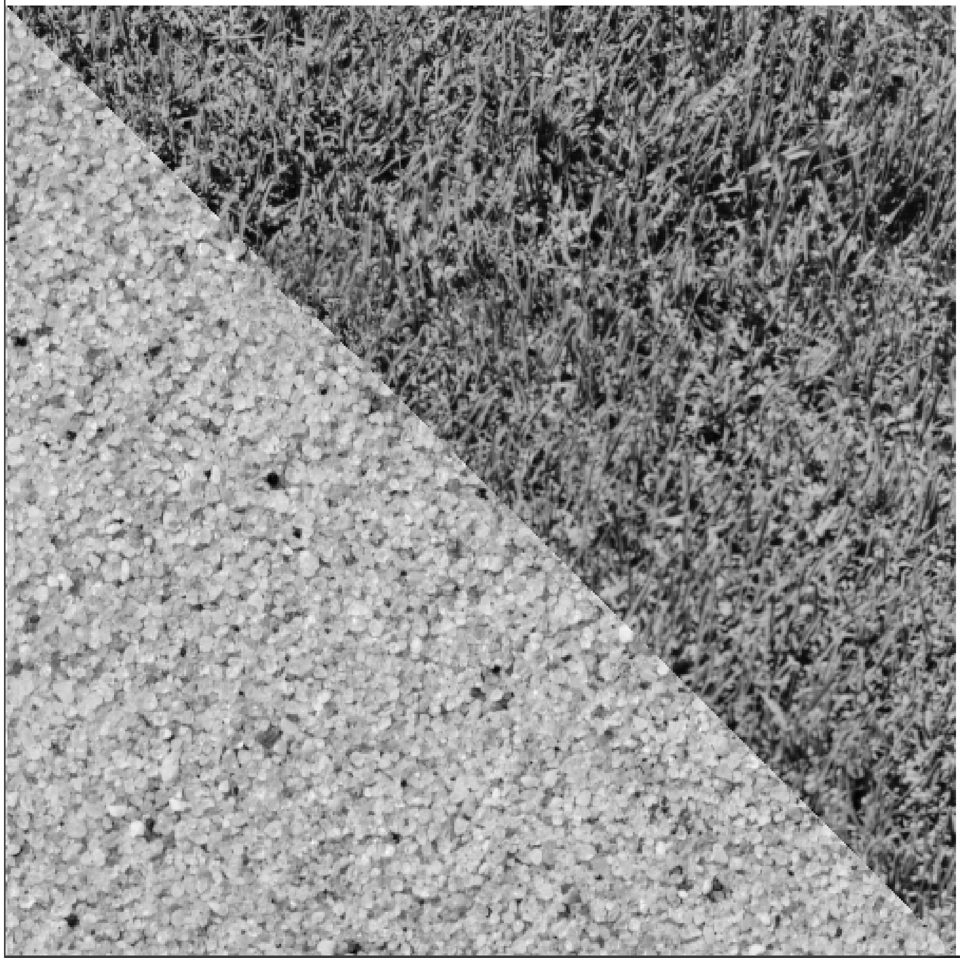
3



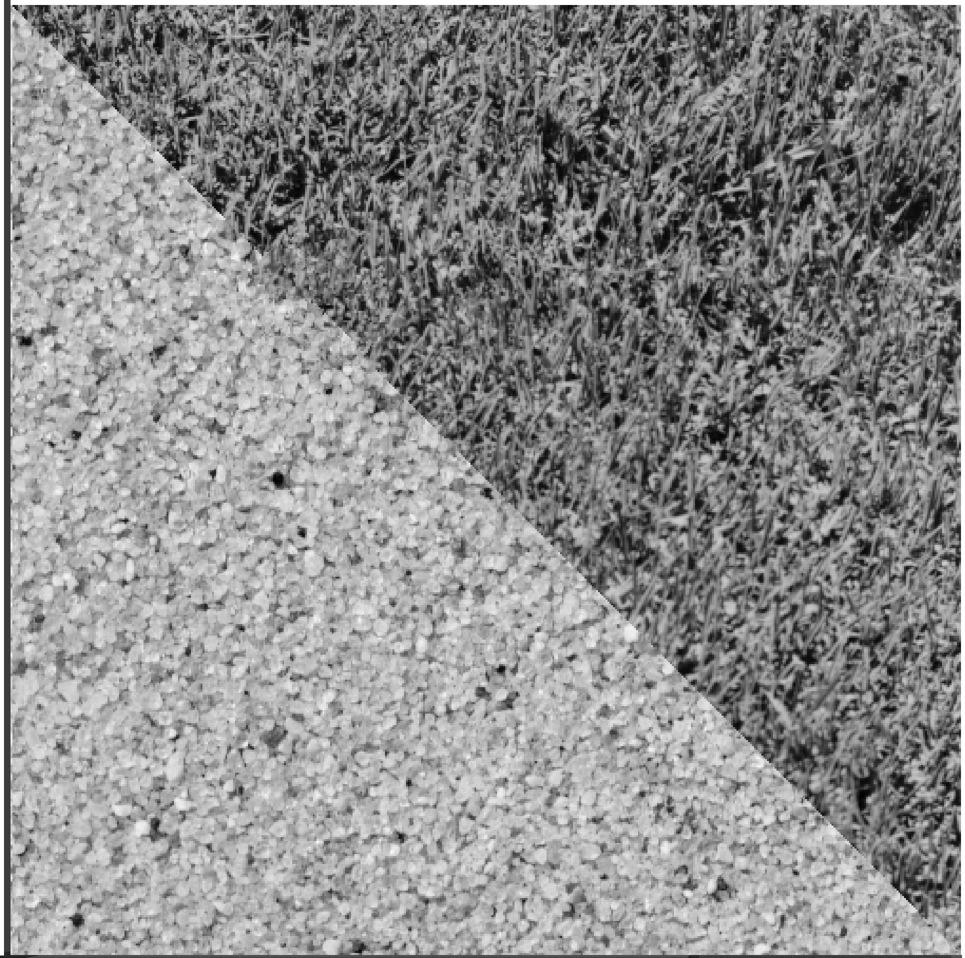


Examples: Quality

2



1



Examples: Comparison

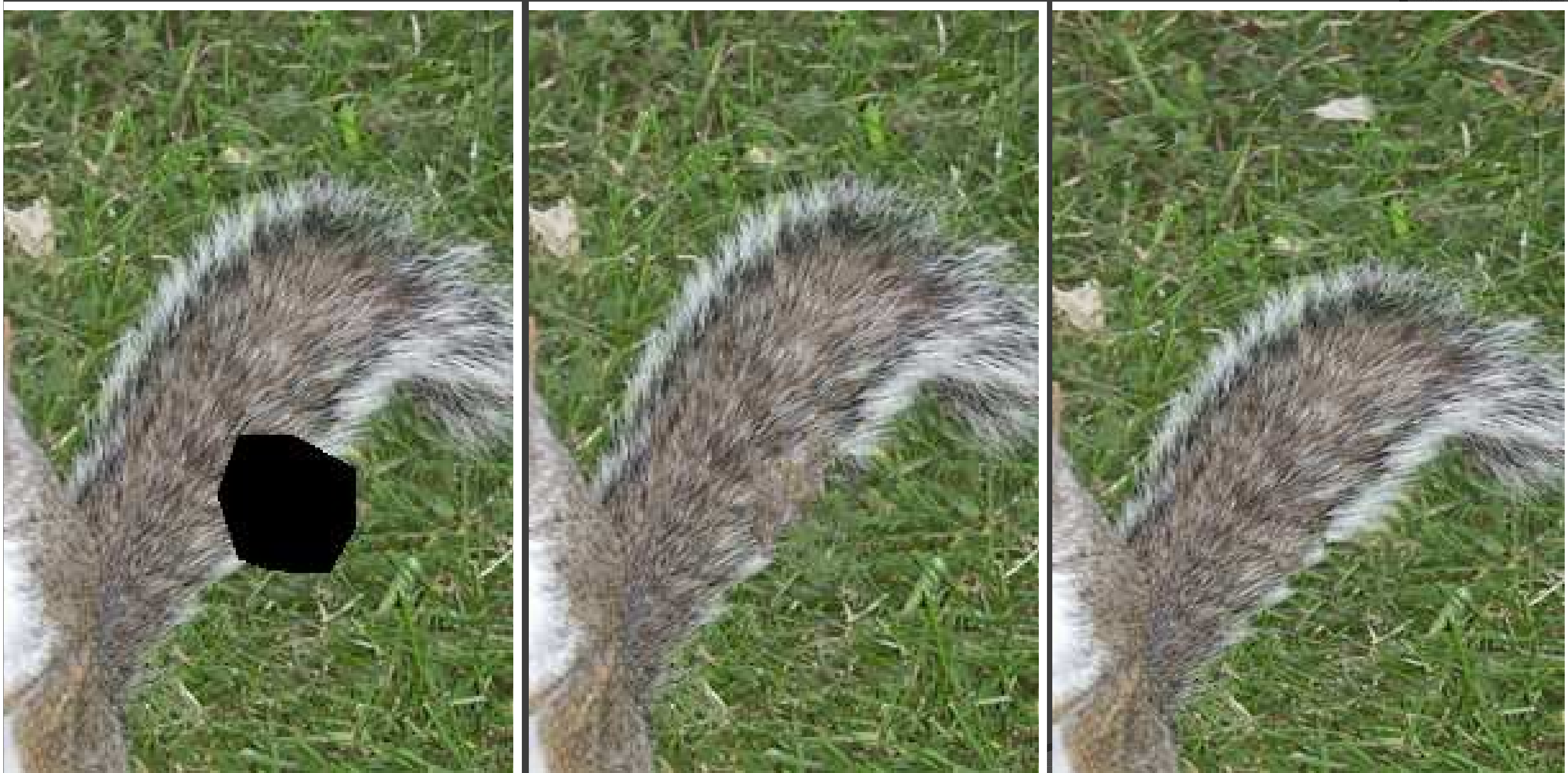
Q	Criminisi	SCIC
1	56%	18%
2	36%	18%
3	8%	64%
Mean Score	1.52	2.46

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