Recap.

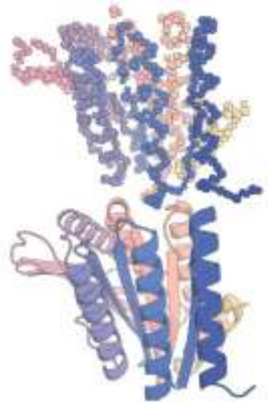
What do these scenarios have in common?

- There are **multiple** or infinite predictions to one input.
- Some predictions are more "**plausible**" than some others. **(High-Quality)**
- Training data may contain **no exact solution**. **(Diversity)**
- Predictions may be **more complex**, more informative, and higher-dimensional than input.

Chatbot

Image generation

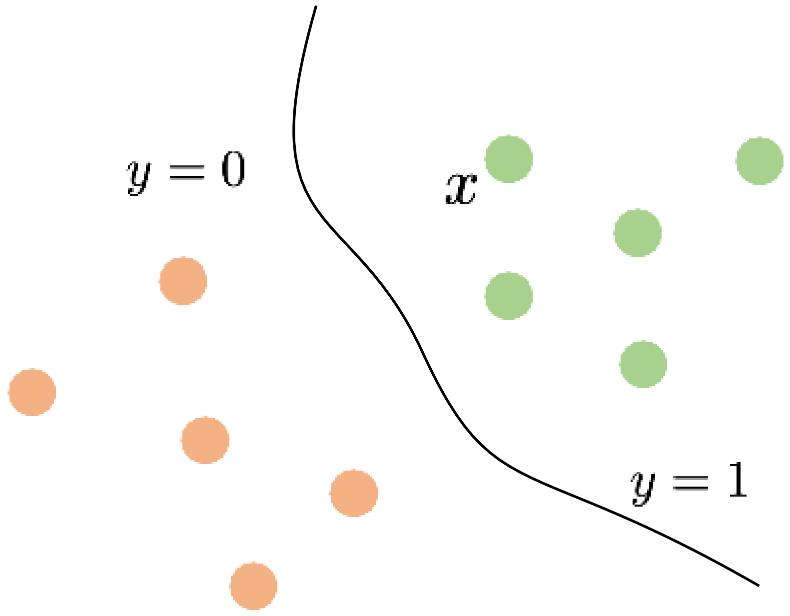
Video generation



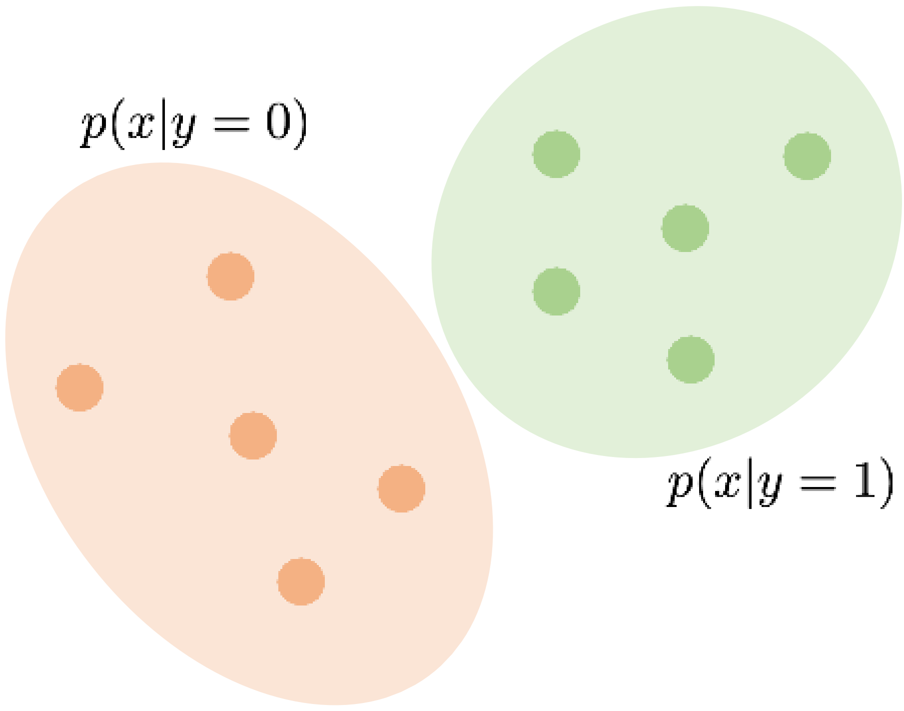
Protein generation

Discriminative vs. Generative models

discriminative $p(y|x)$



generative $p(x|y)$

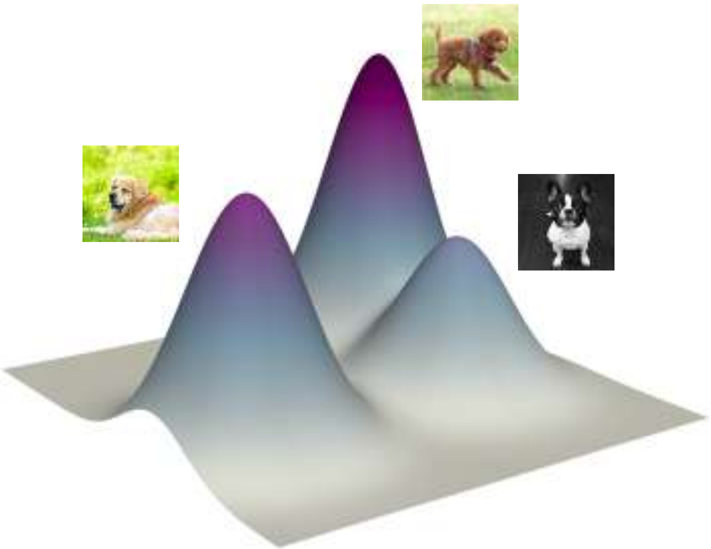


- Generative models can be discriminative: Bayes’ rule
- Can discriminative models be generative?

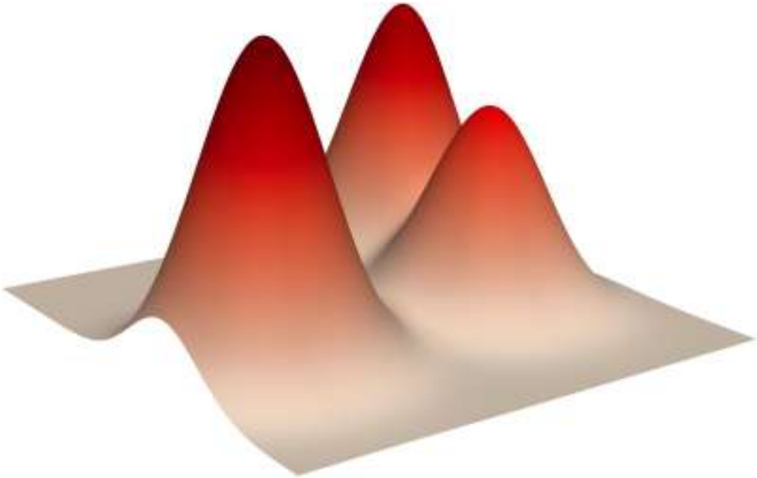
Generative models w/ probabilistic modeling

data

distribution
of data



estimated
distribution
of data

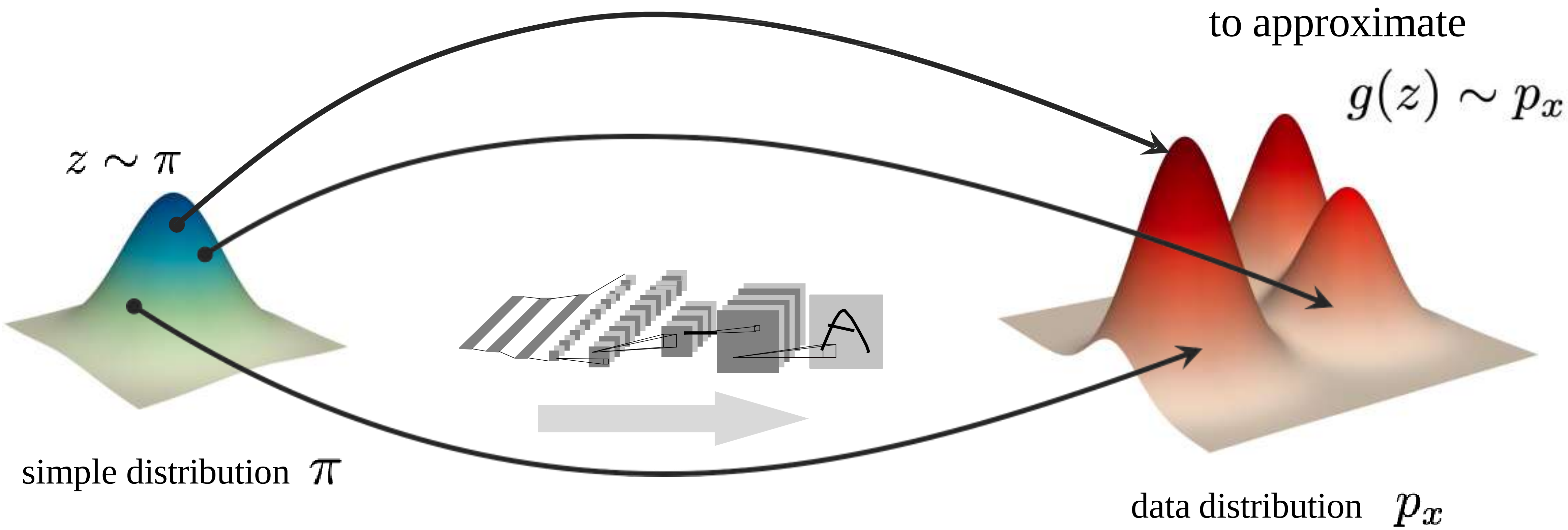


- Optimize a loss function

$\mathcal{L}($  ,  $)$

Learning to represent probability distributions

- From simple to complex distributions



Formulating Real-world Problems as Generative Models

- Generative models are about $p(x|y)$

What can be *y*?

- condition
- constraint
- labels
- attributes

- more abstract
- less informative

What can be *x*?

- "data"
- samples
- observations
- measurements

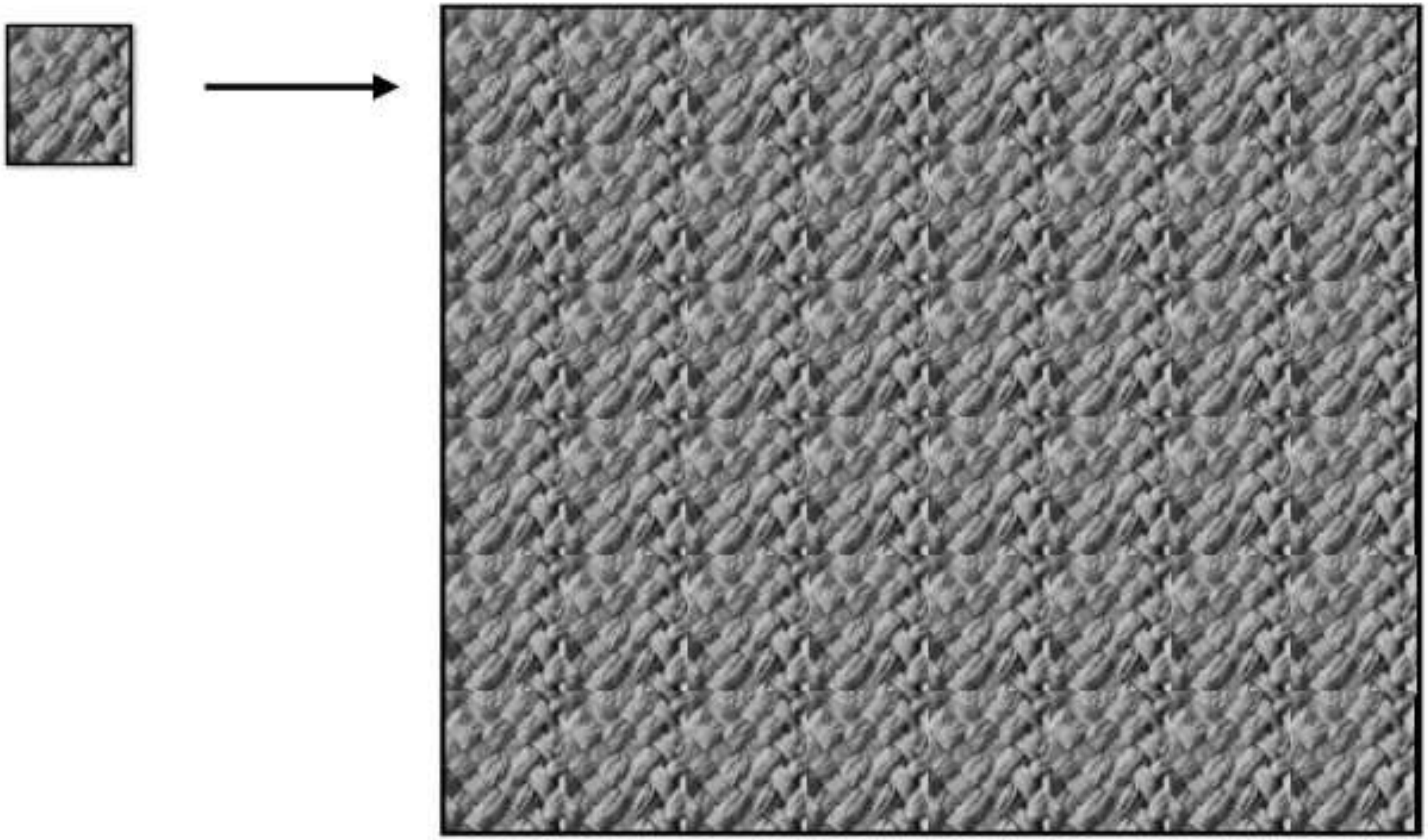
- more concrete
- more informative

Formulating Real-world Problems as Generative Models

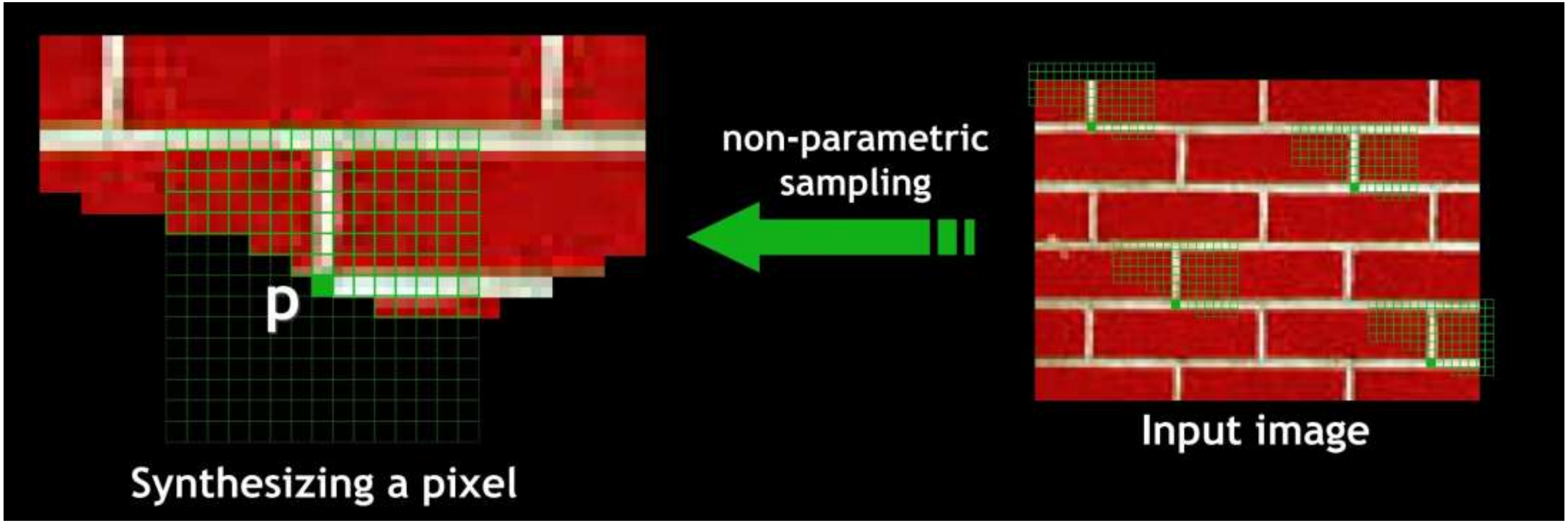
- Generative models are about $p(x|y)$
- Many problems can be formulated as generative models
- What's *x*? What's *y*?
- How to represent *x*, *y*, and their dependence?

Image Manipulation with Example

How to Generate Images?

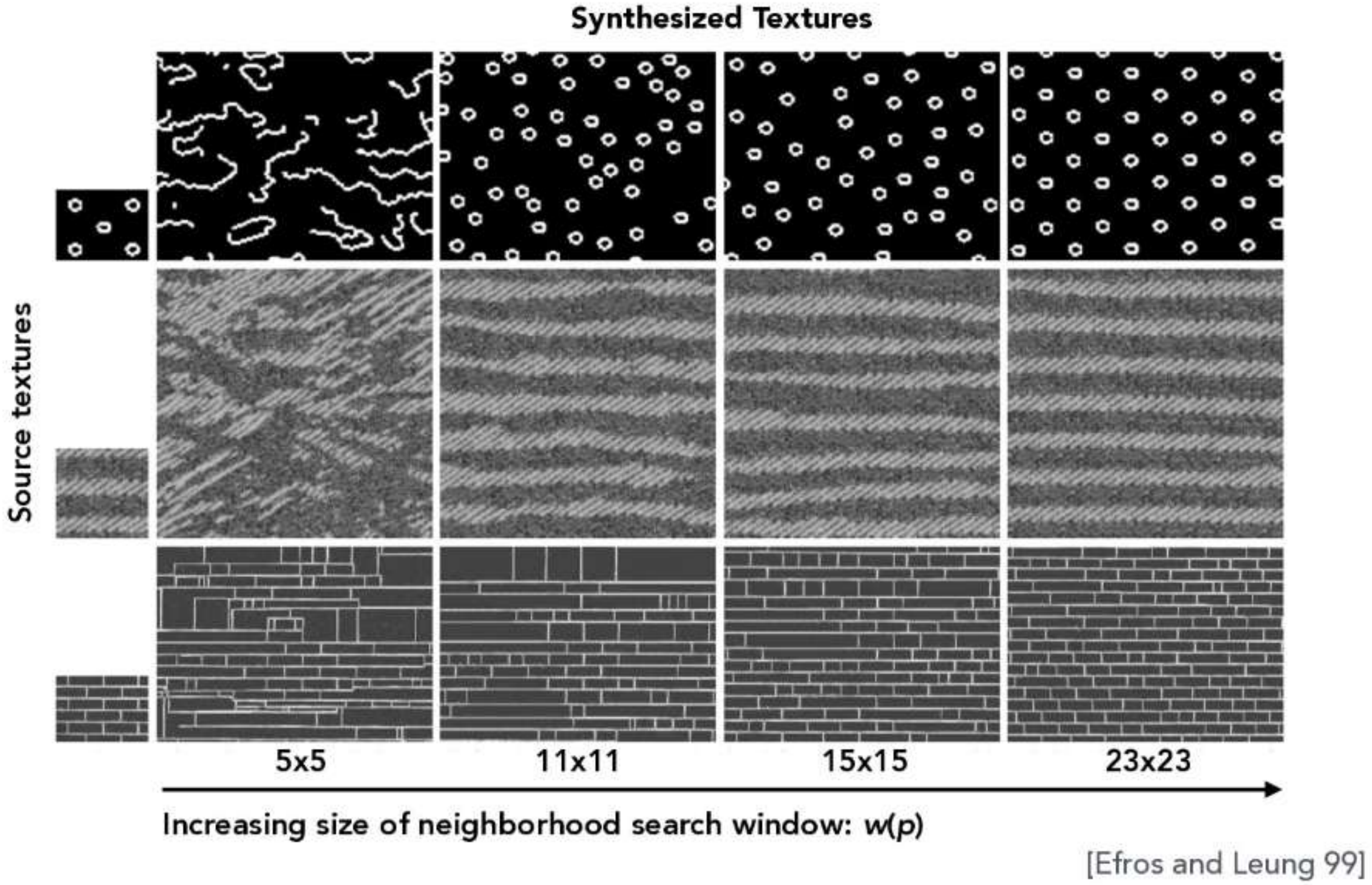


Generative Models before the “GenAI” Era



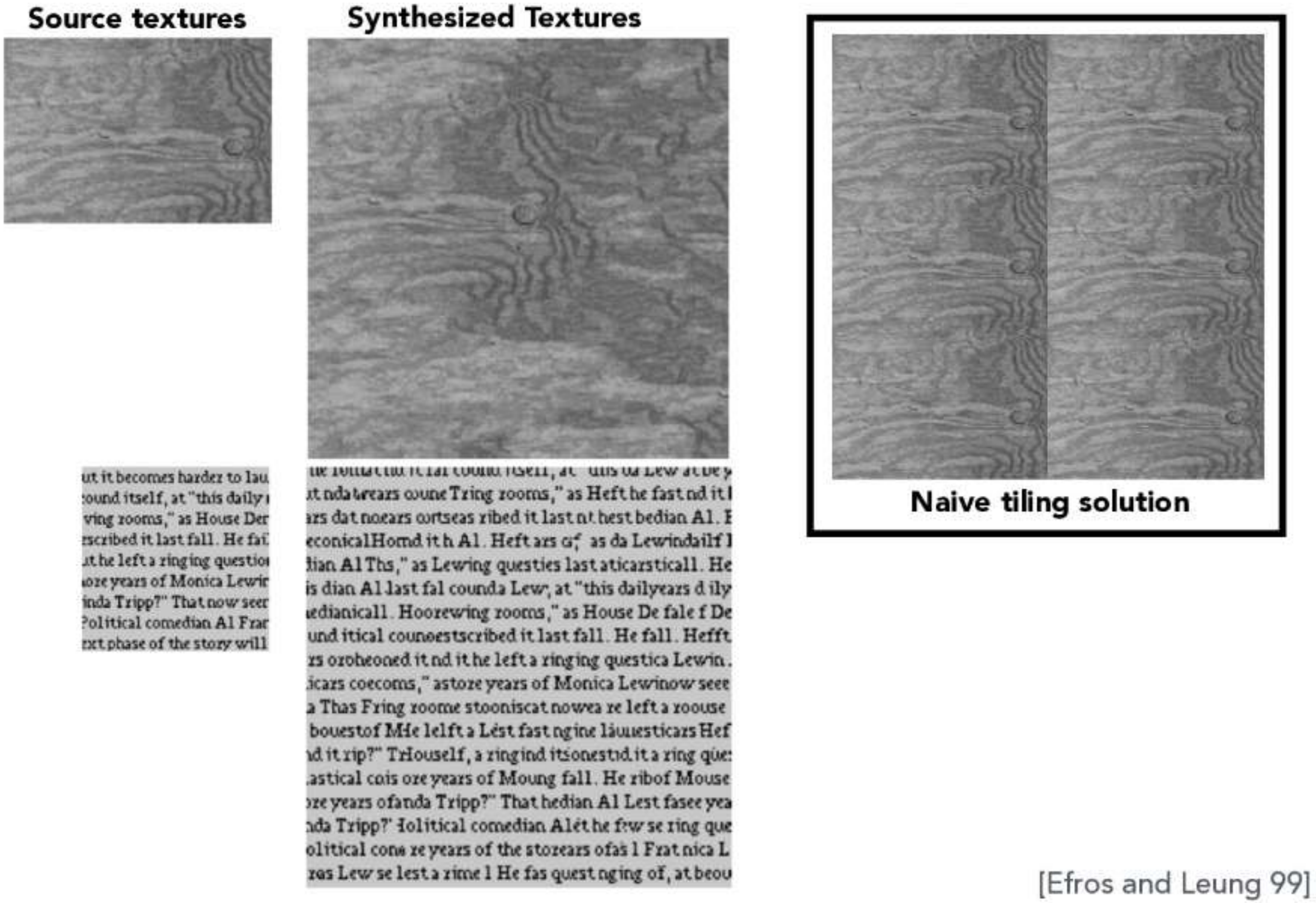
Texture Synthesis by Non-parametric Sampling, ICCV 1999

Generative Models before the "GenAI" Era



Texture Synthesis by Non-parametric Sampling, ICCV 1999

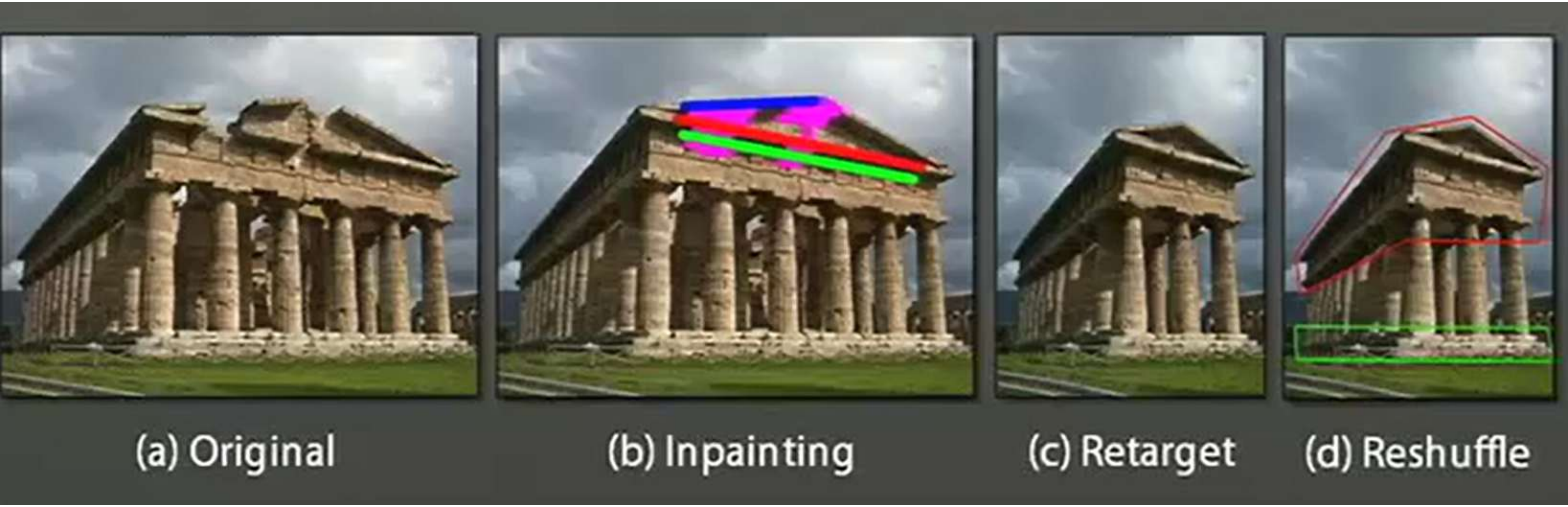
Generative Models before the “GenAI” Era



Texture Synthesis by Non-parametric Sampling, ICCV 1999

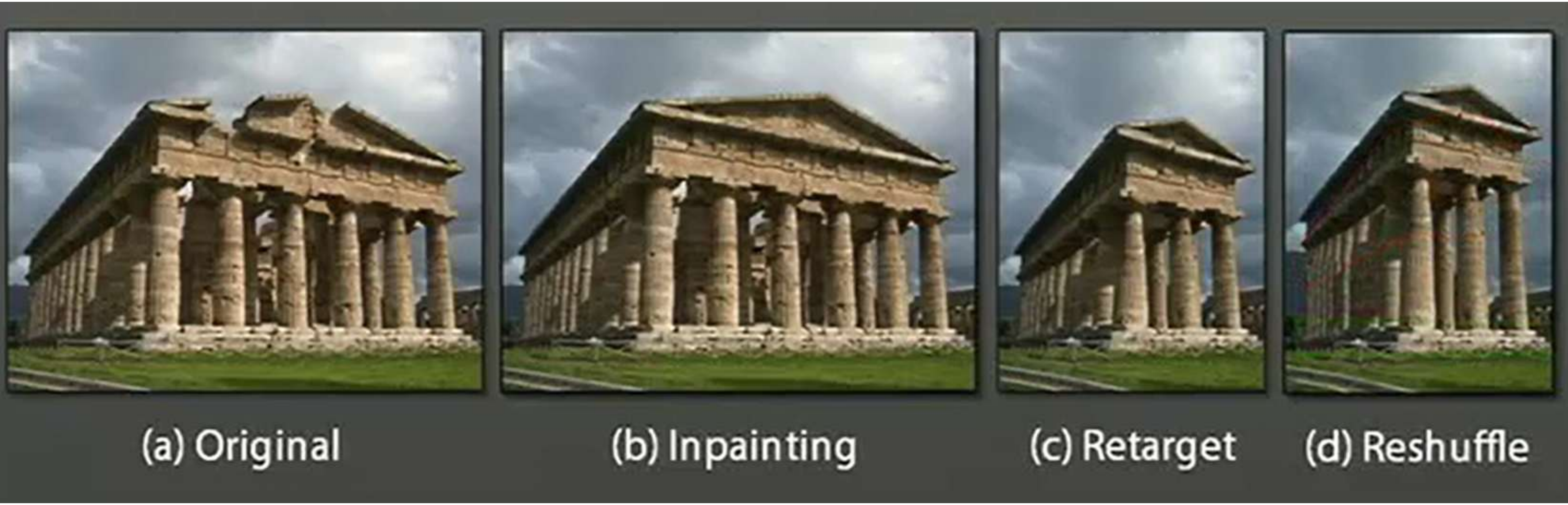
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



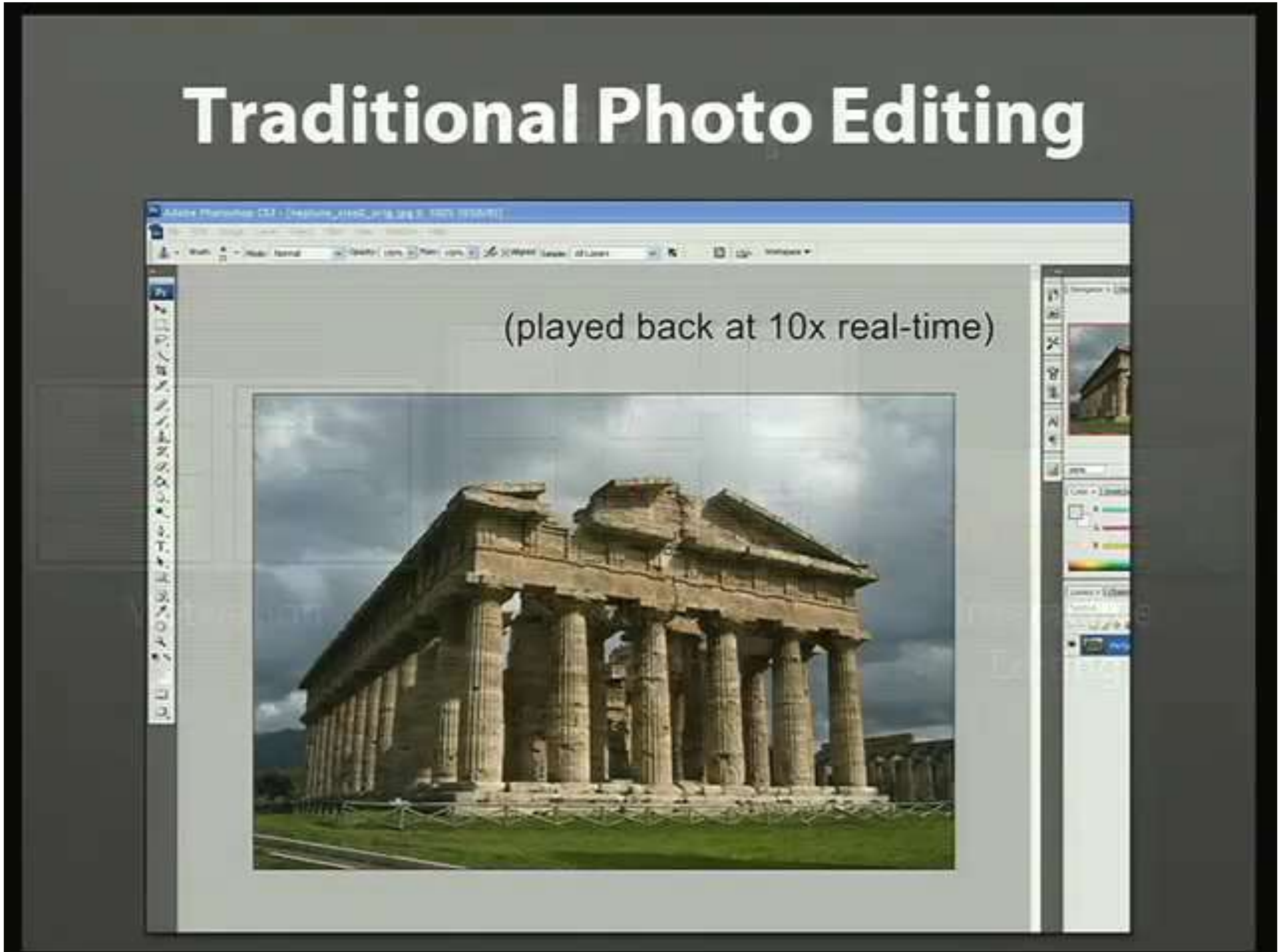
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



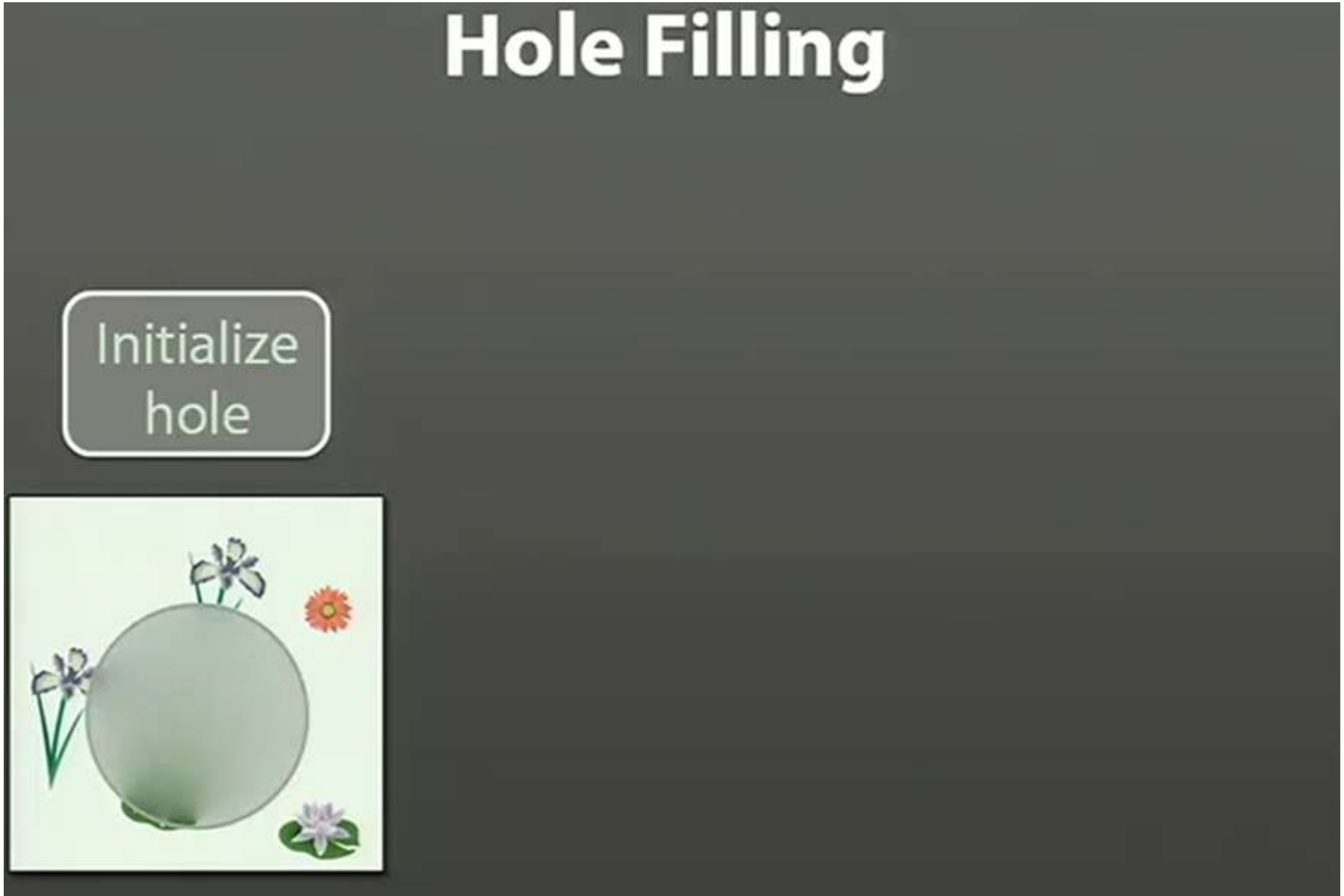
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



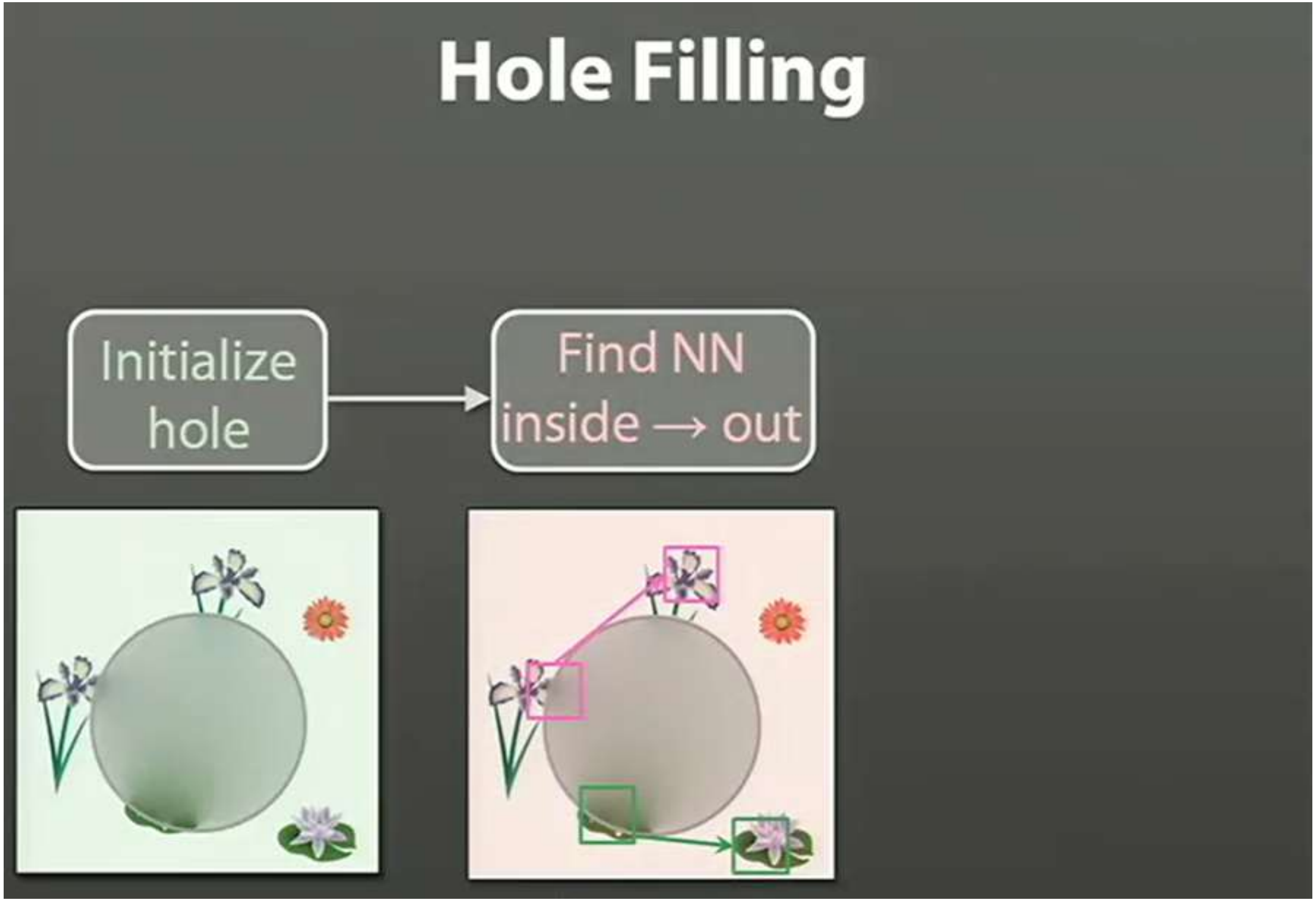
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



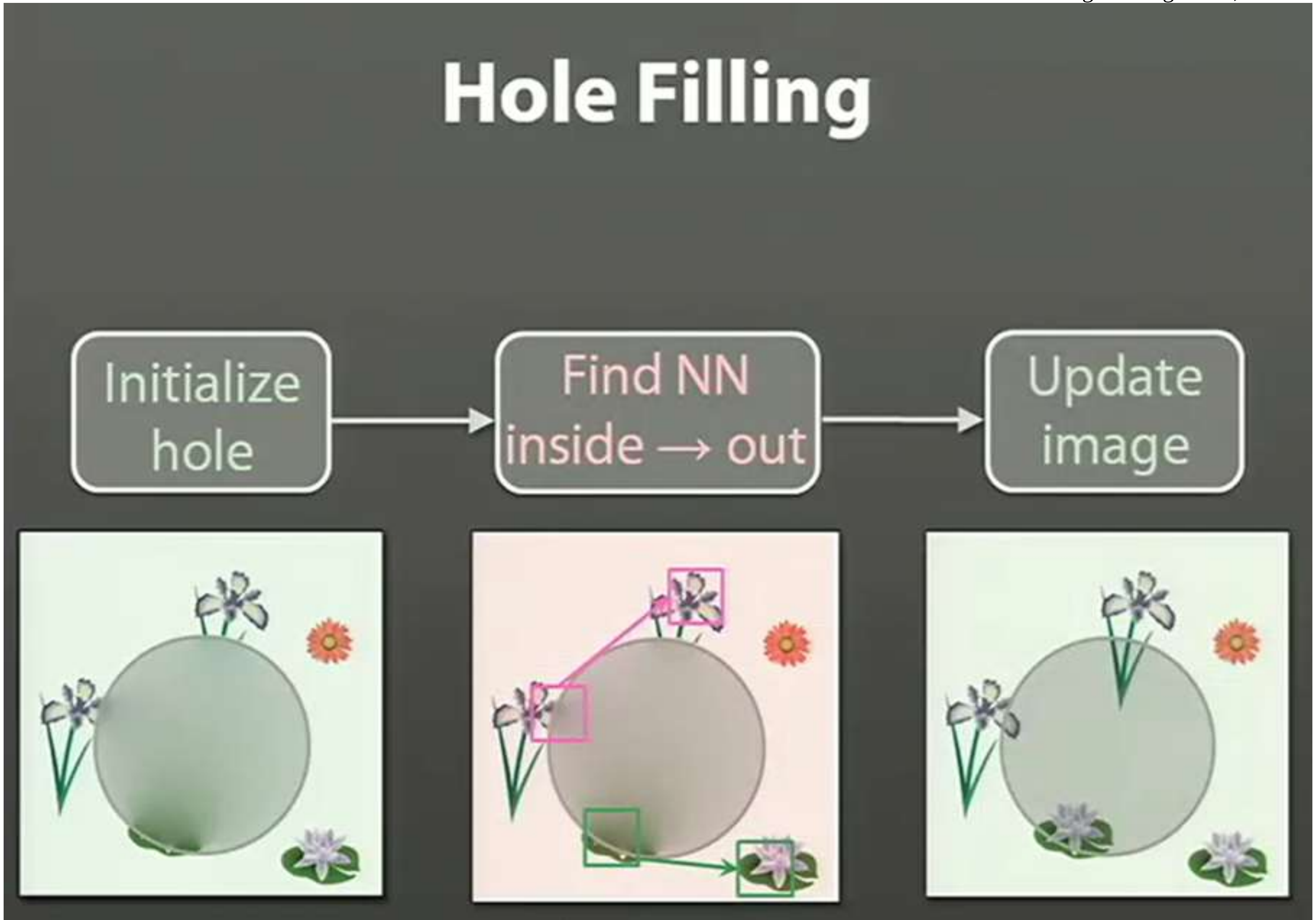
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



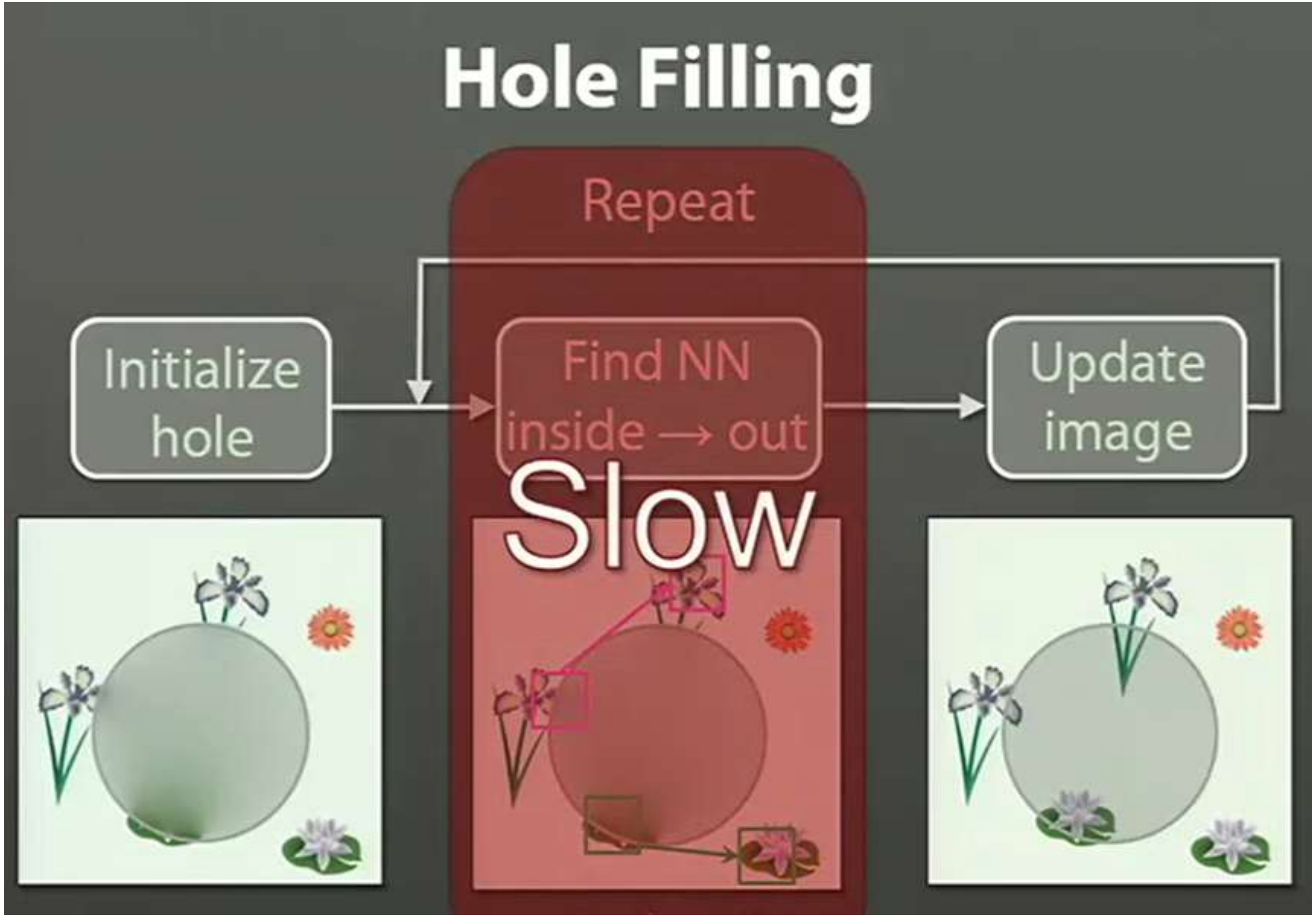
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



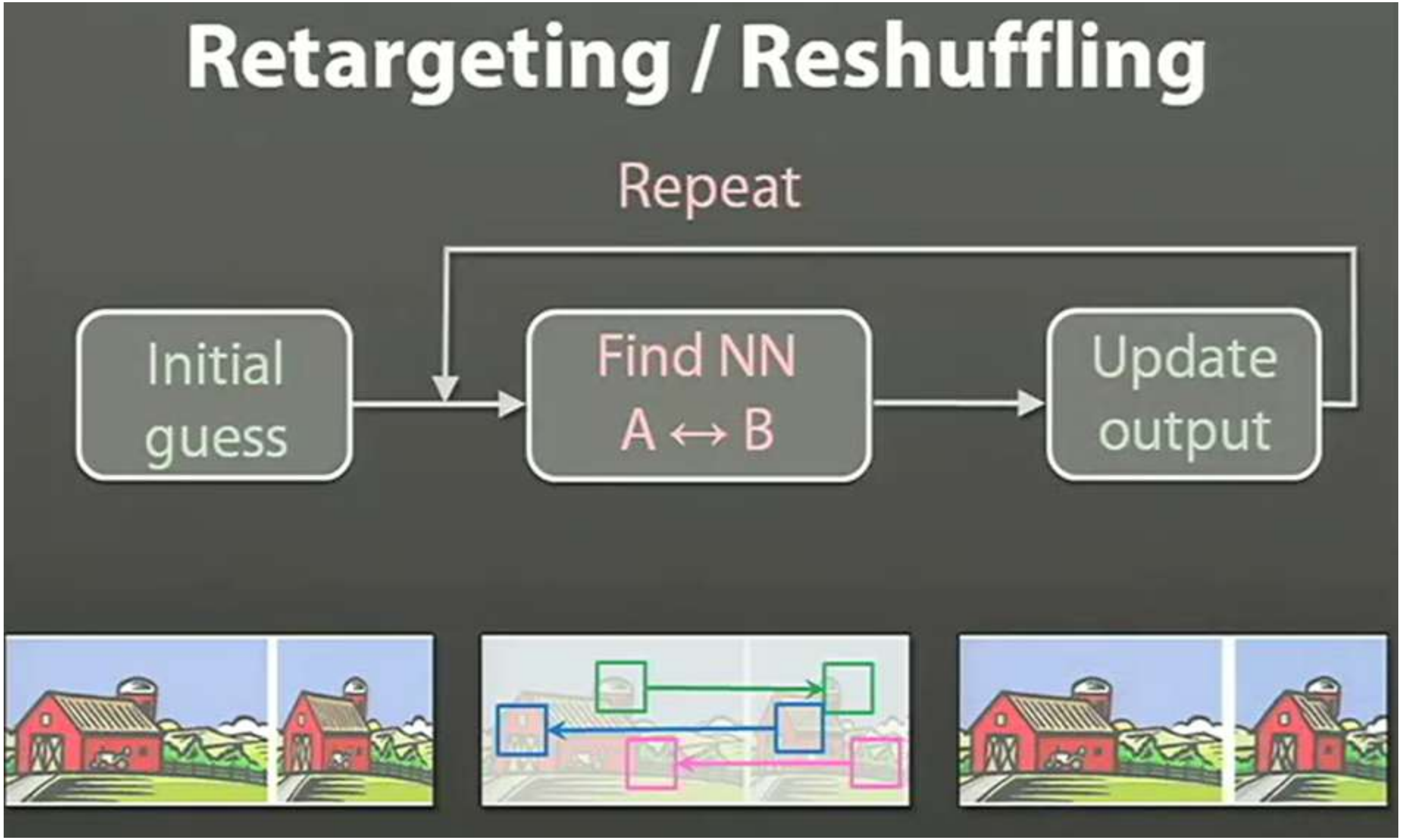
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



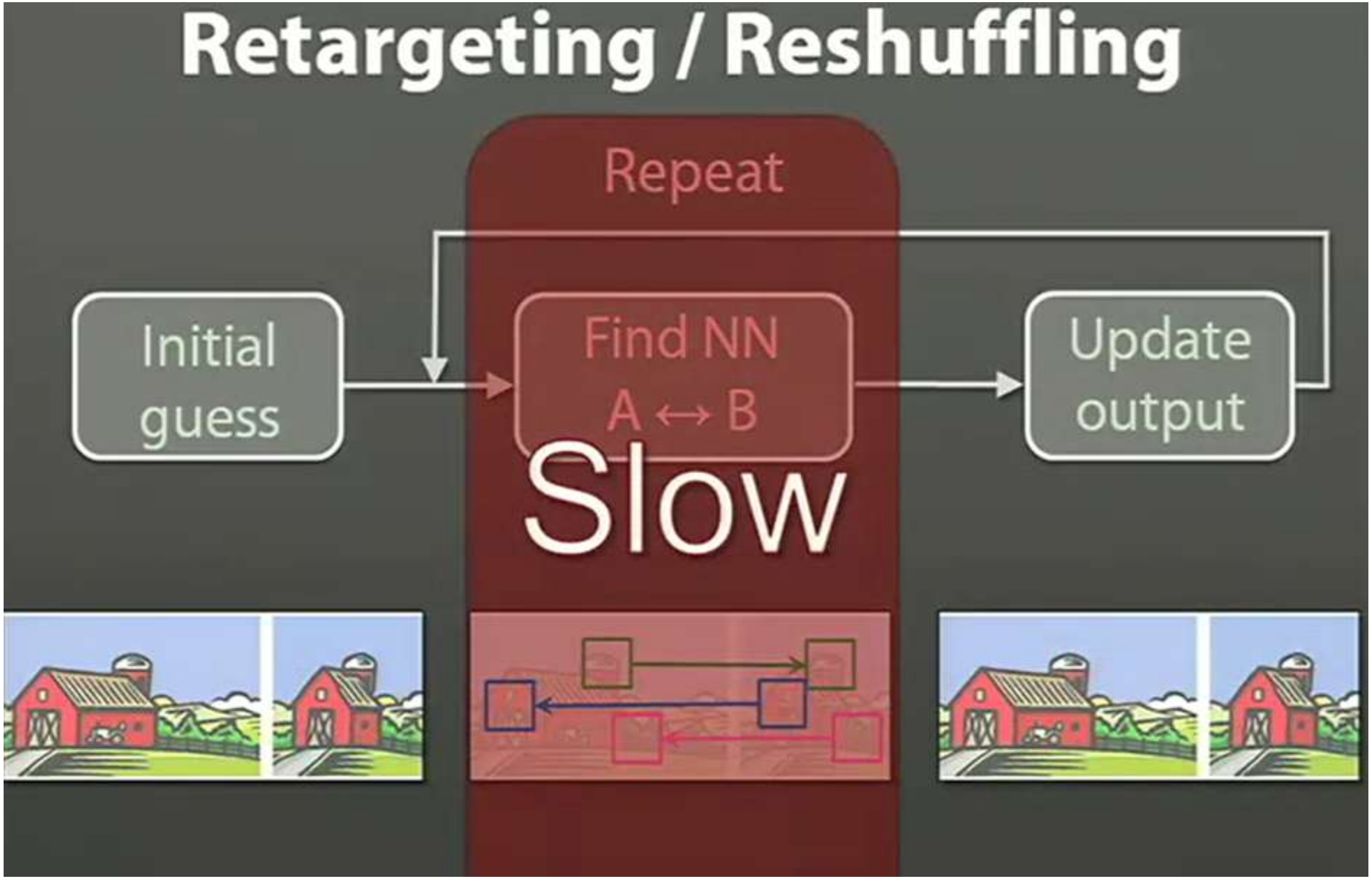
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



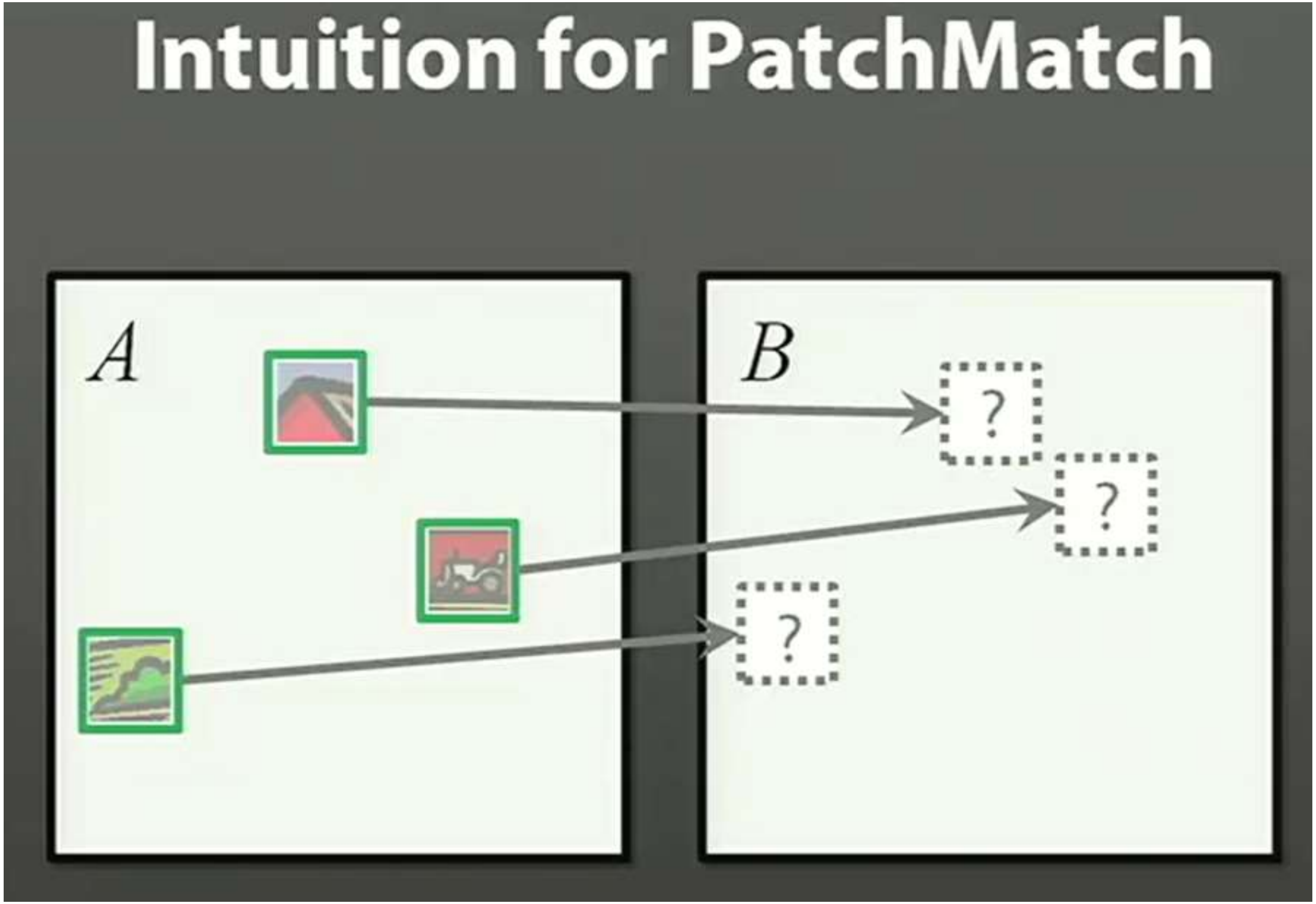
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



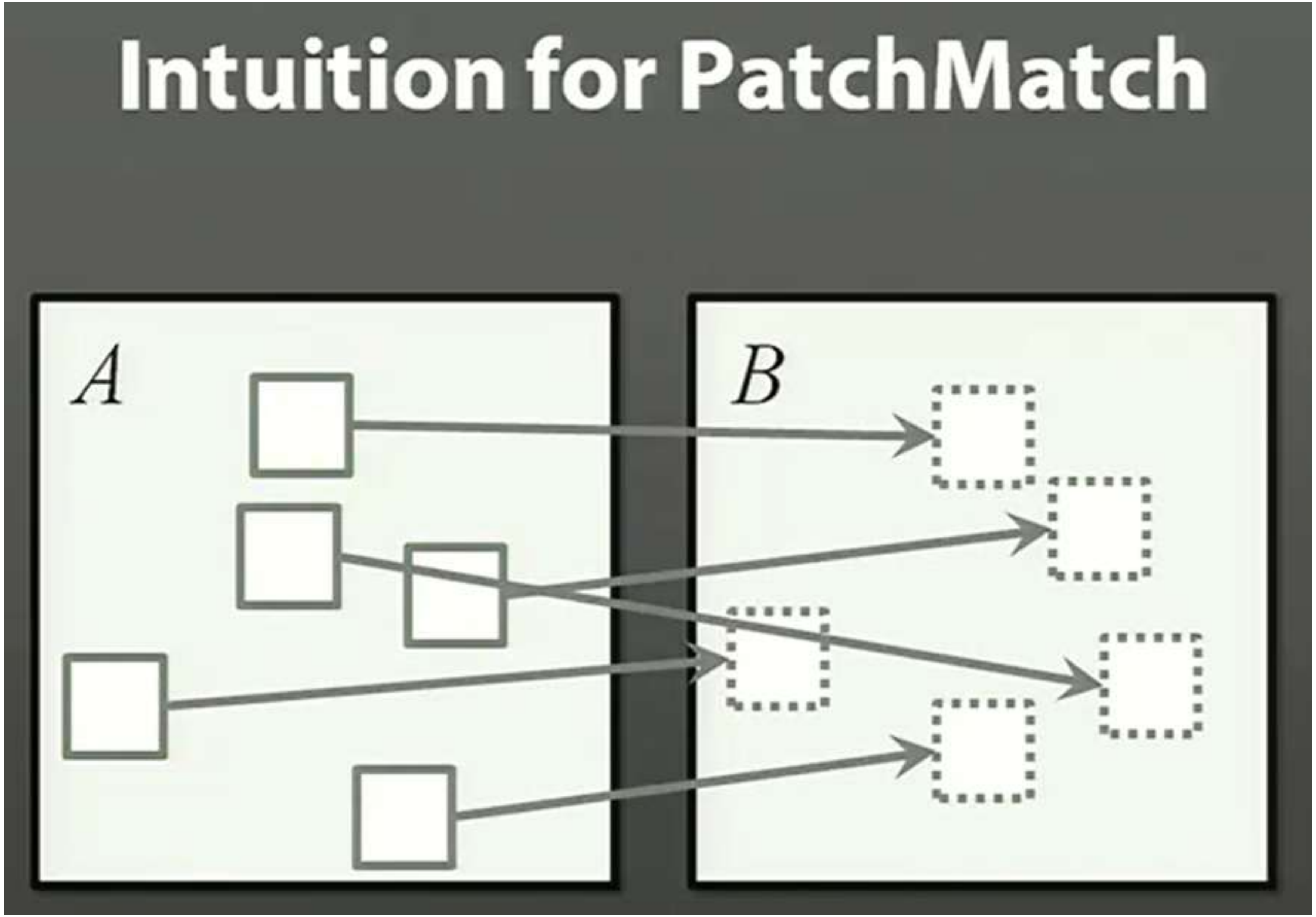
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



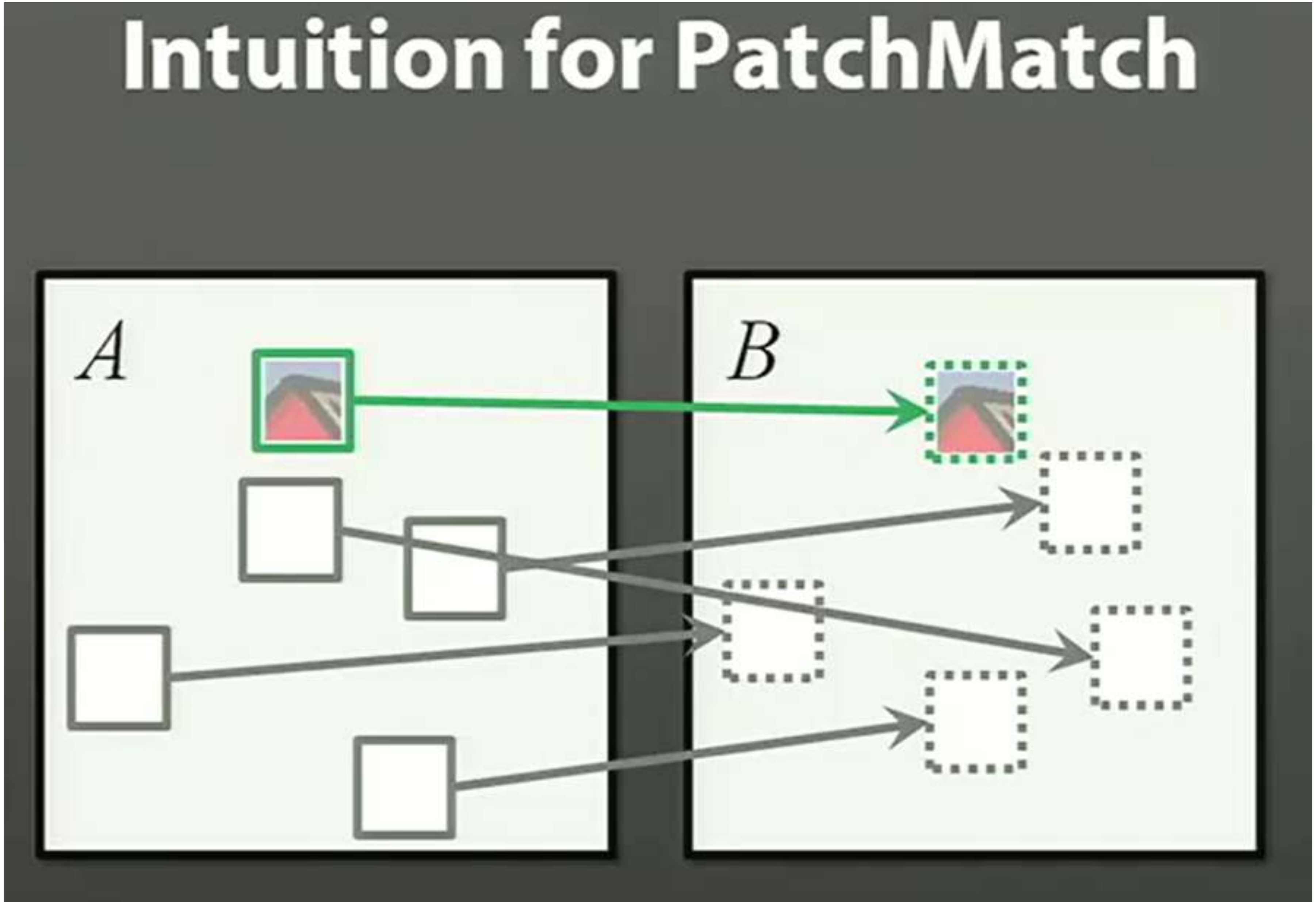
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



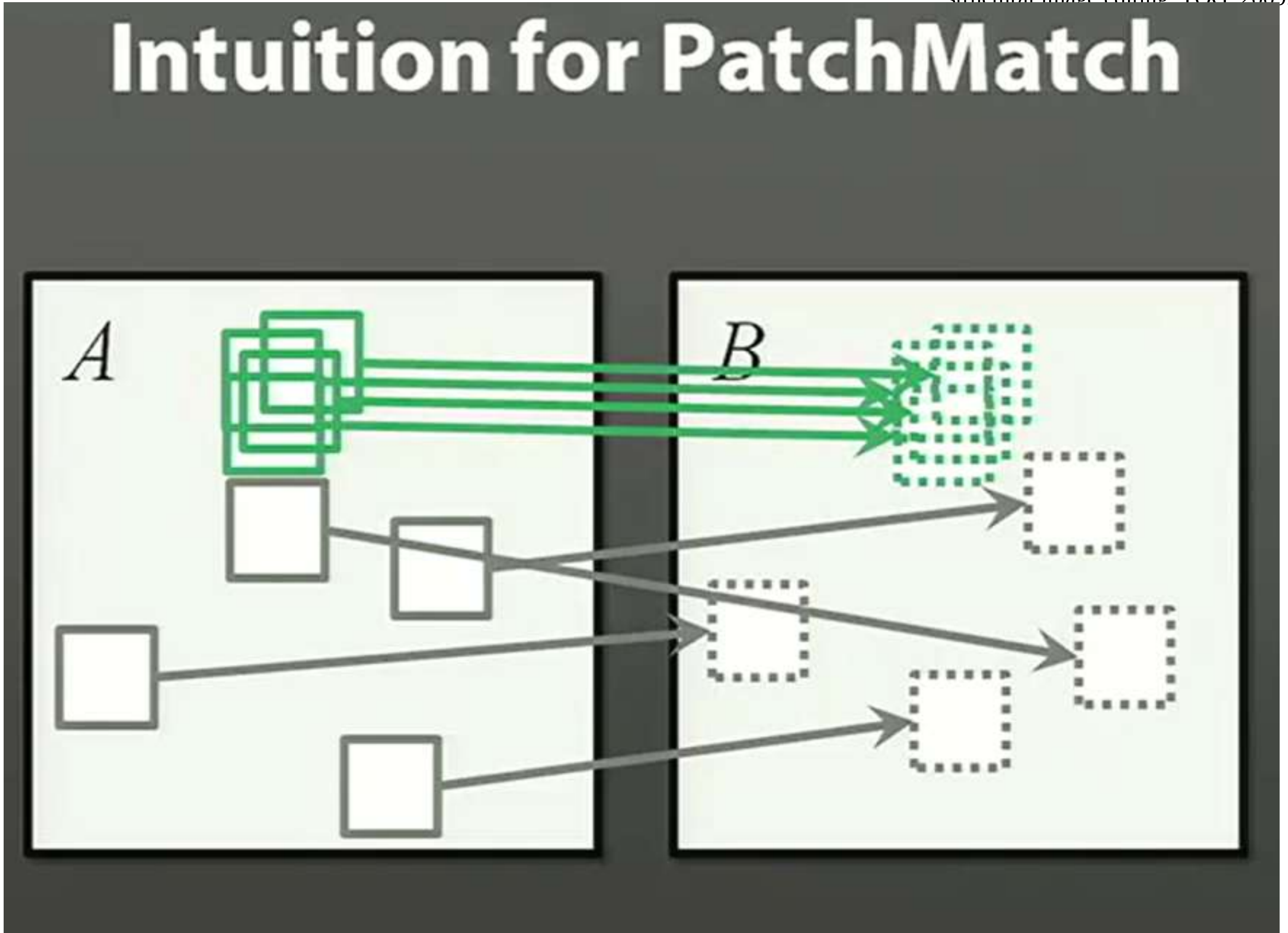
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



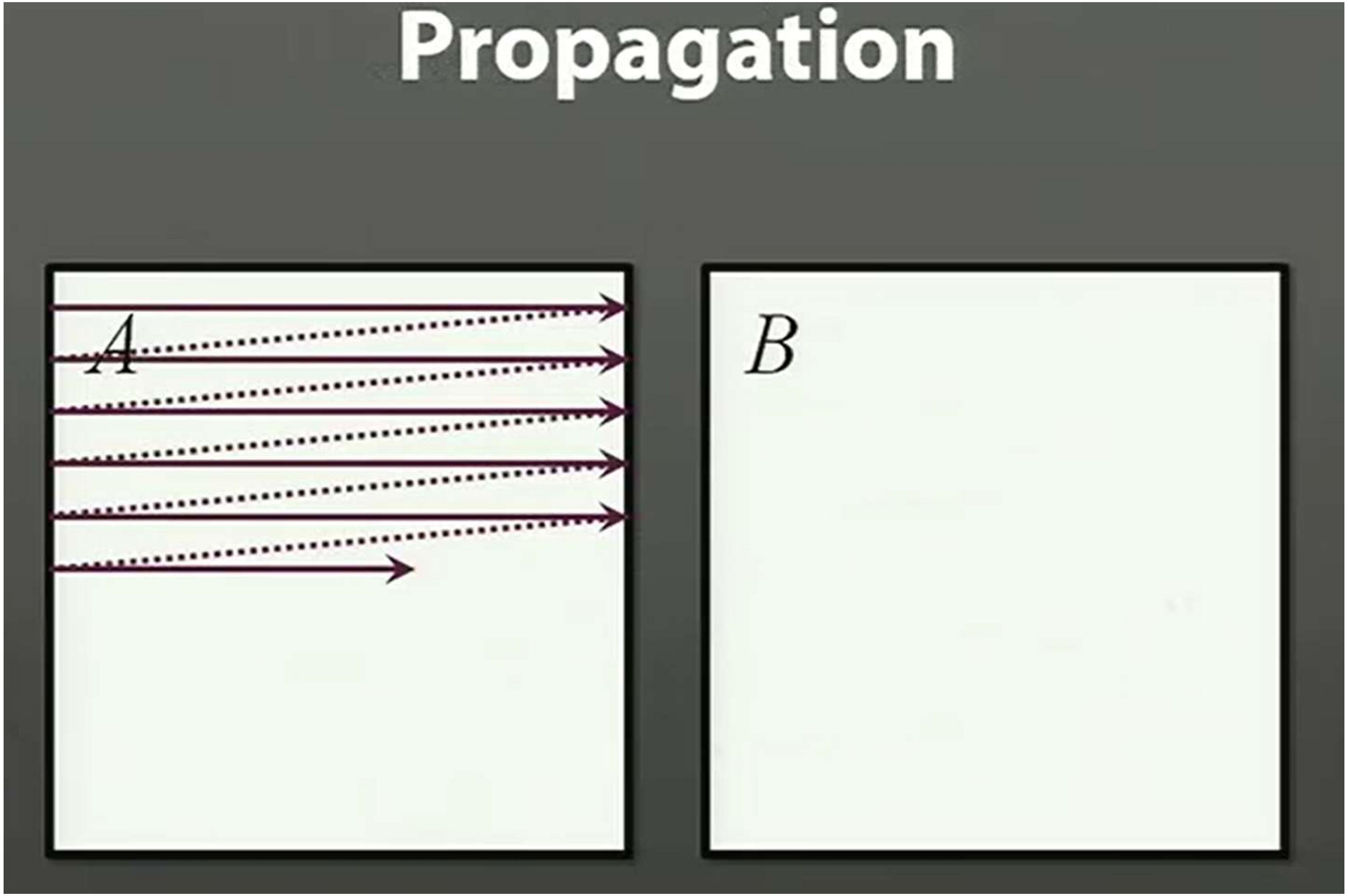
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



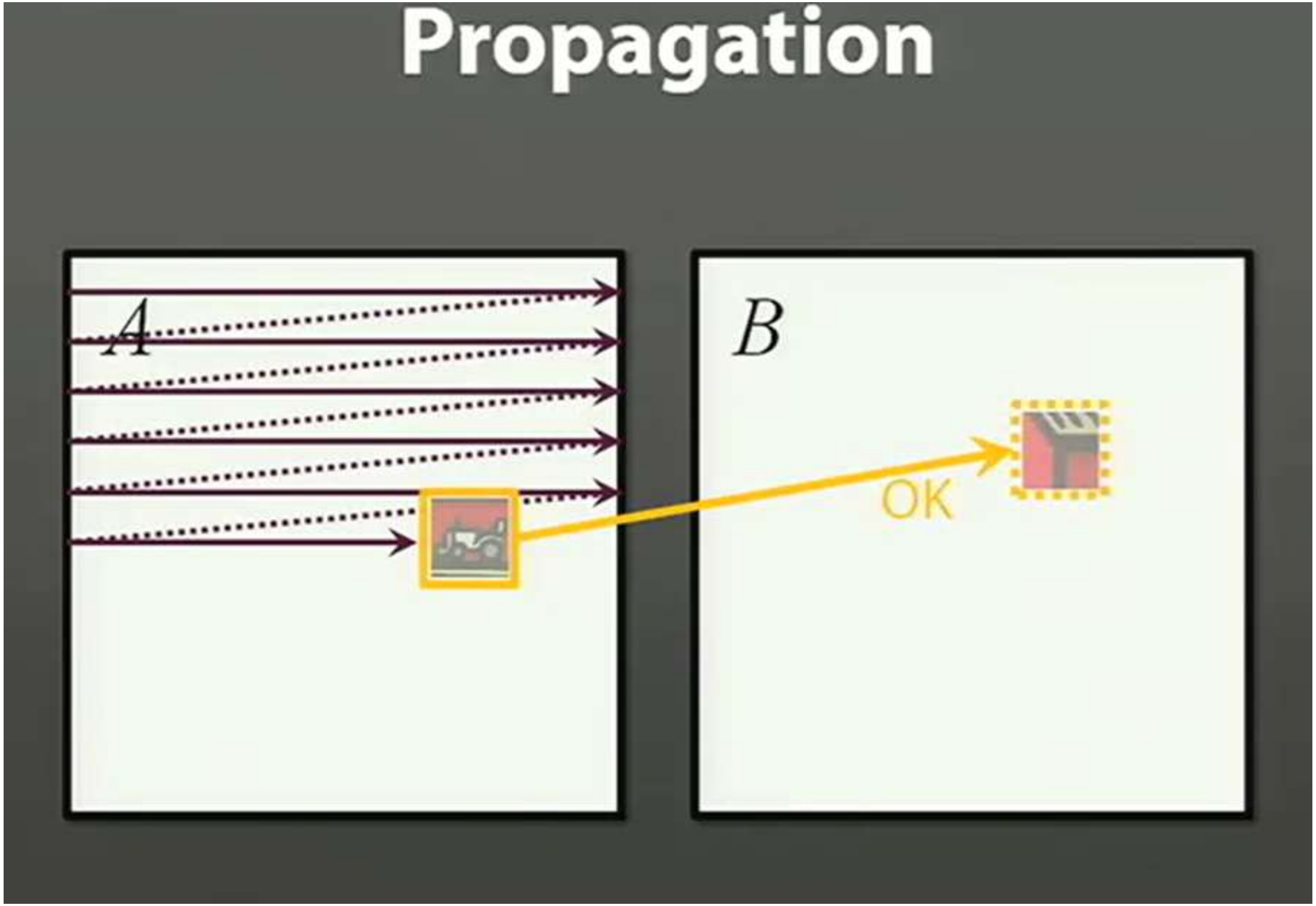
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



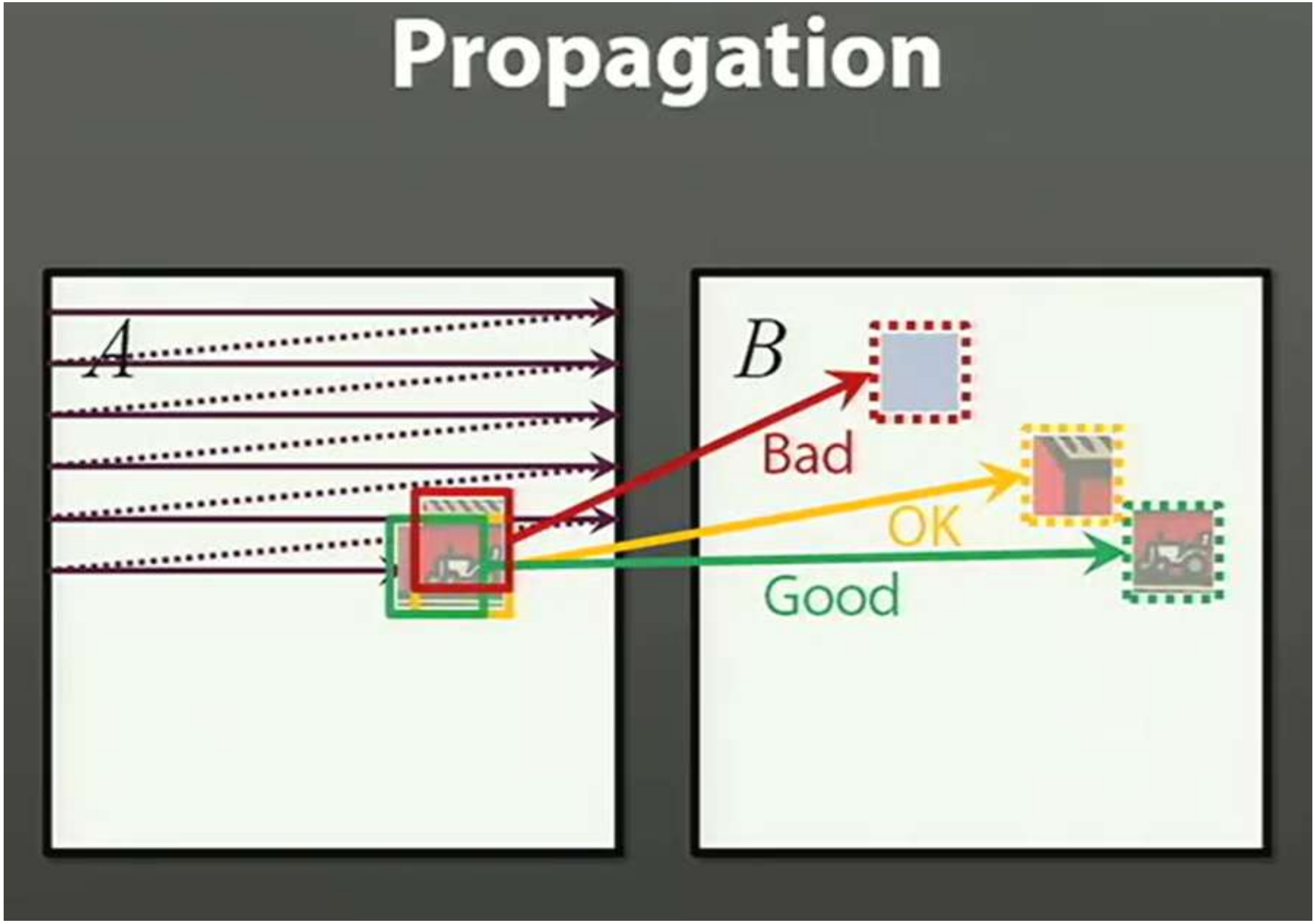
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



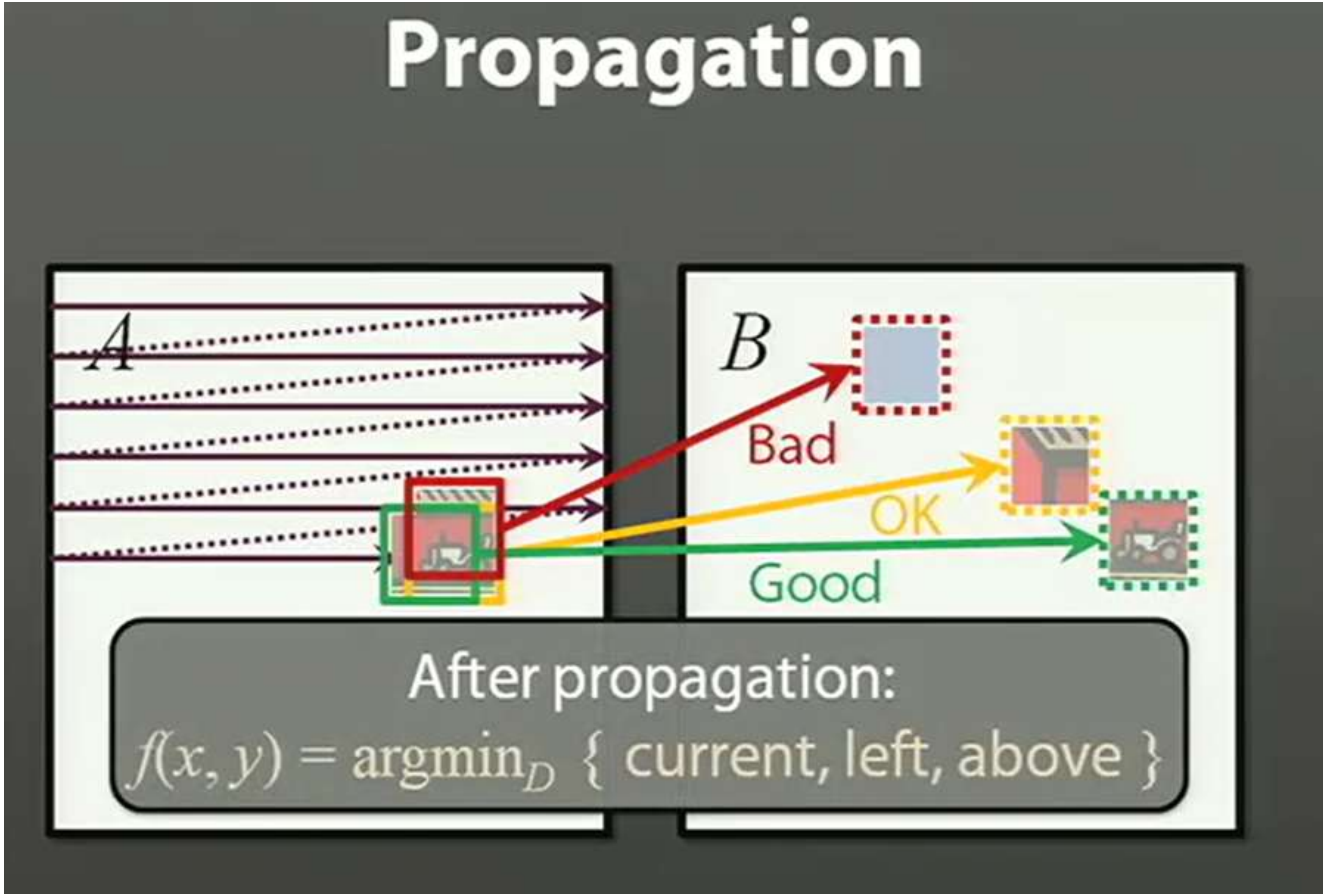
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



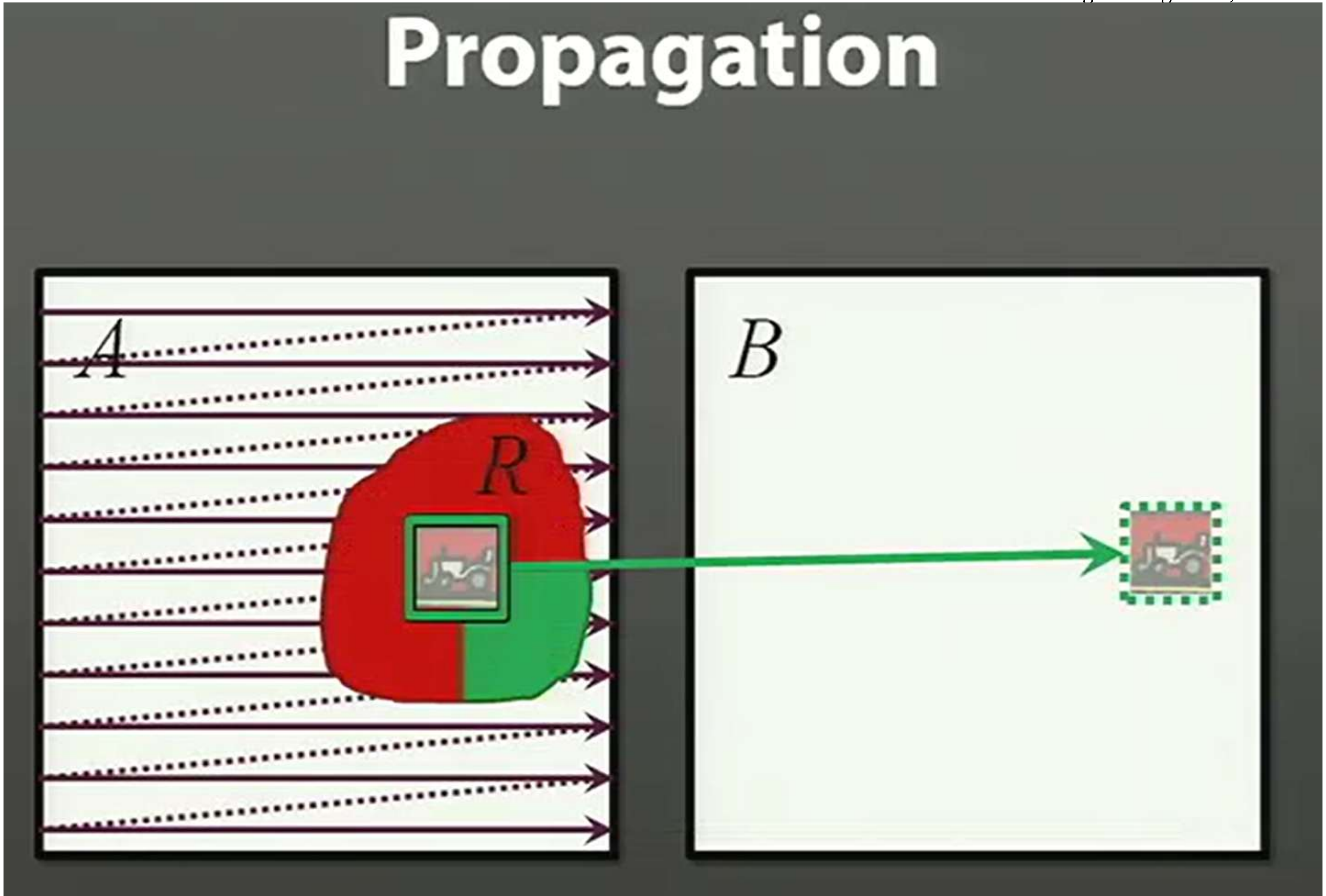
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



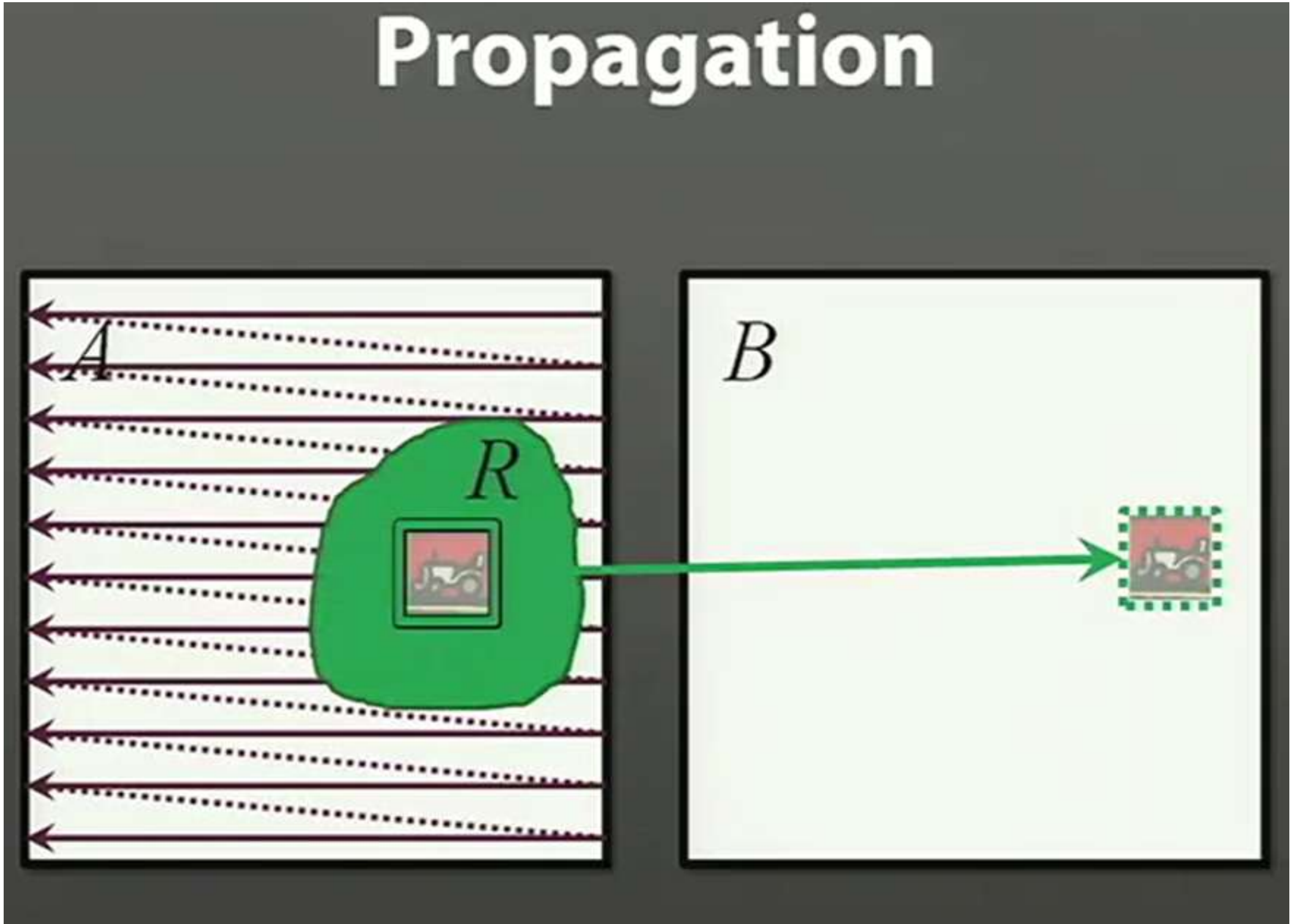
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



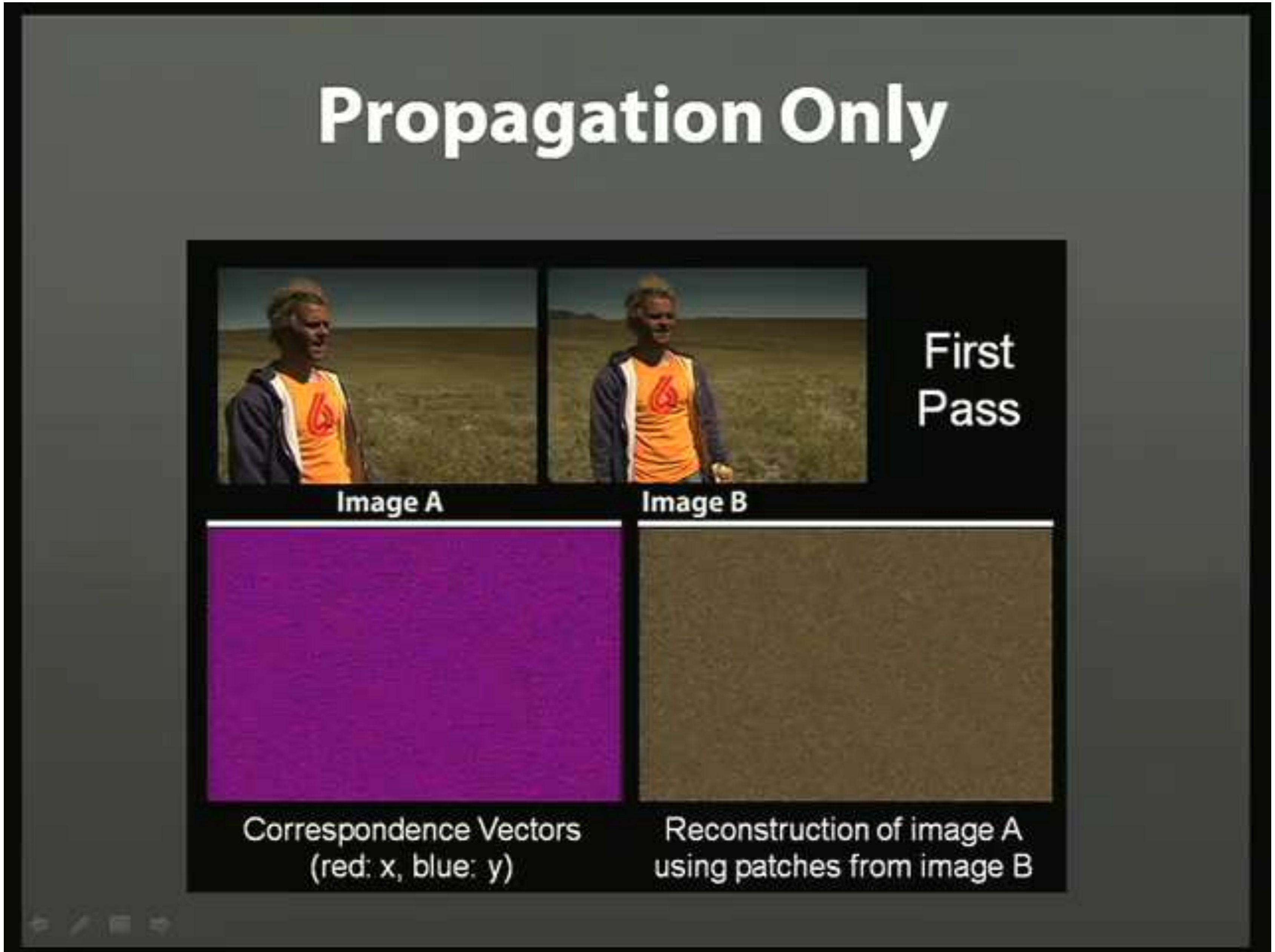
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



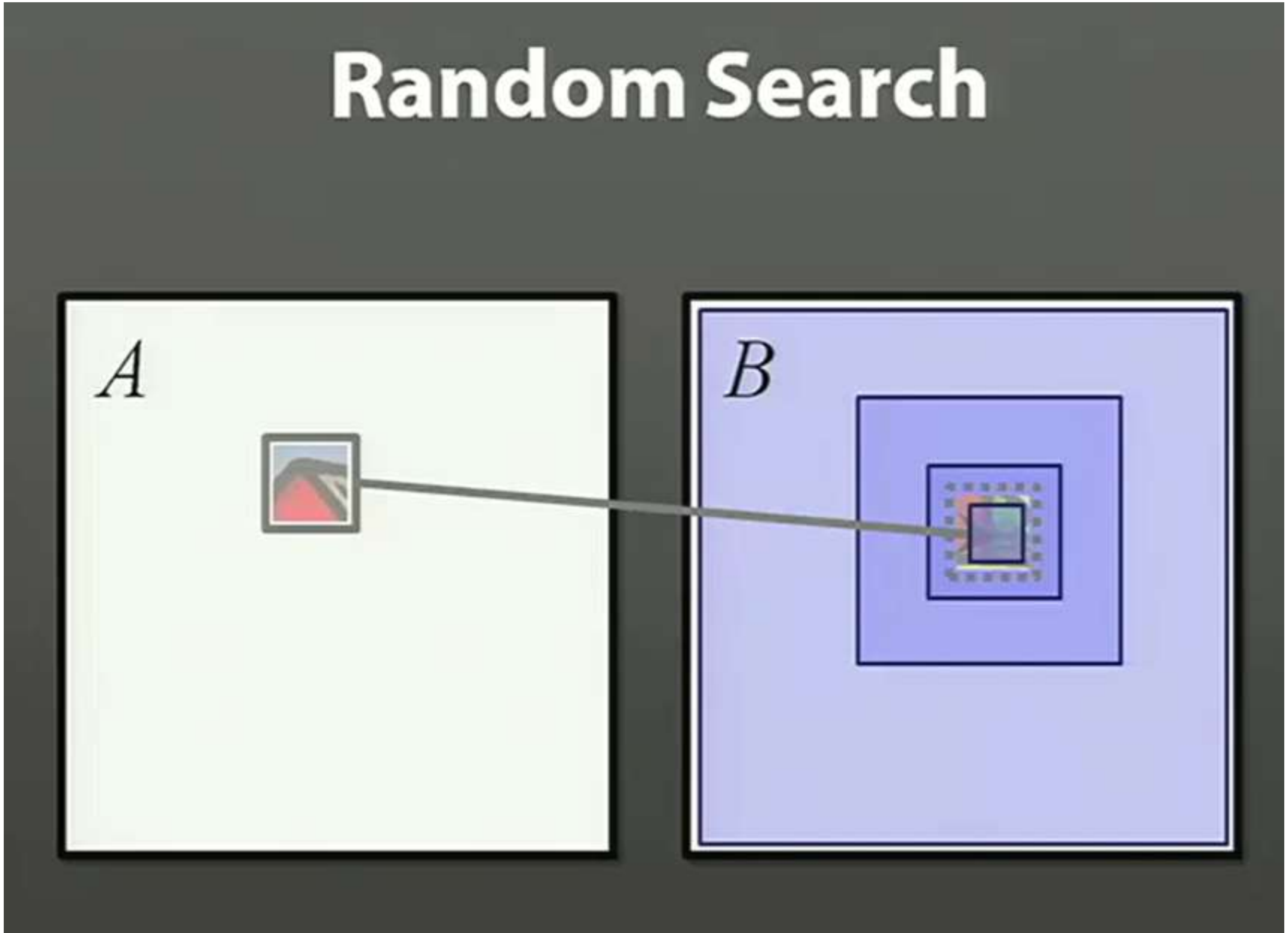
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



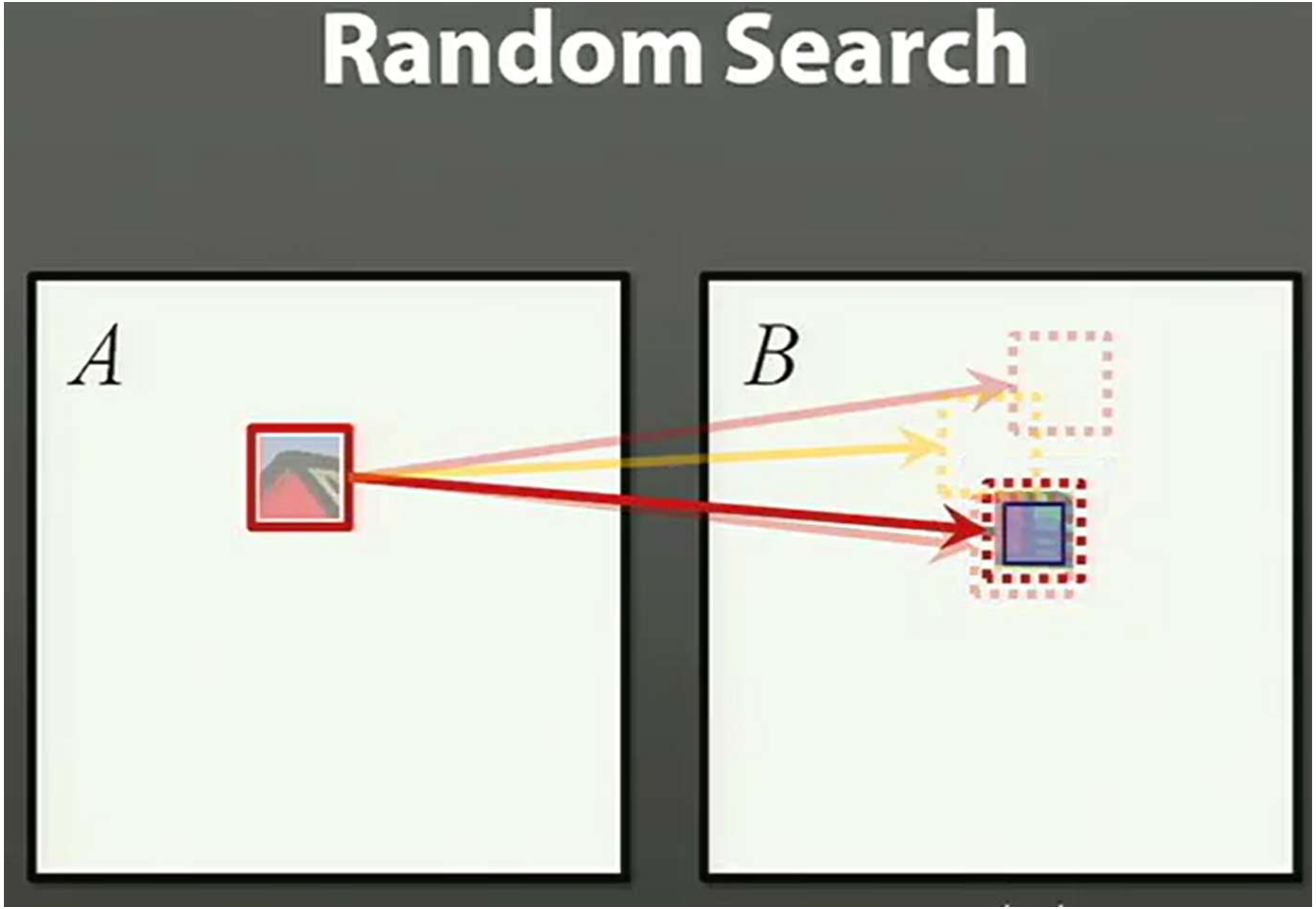
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



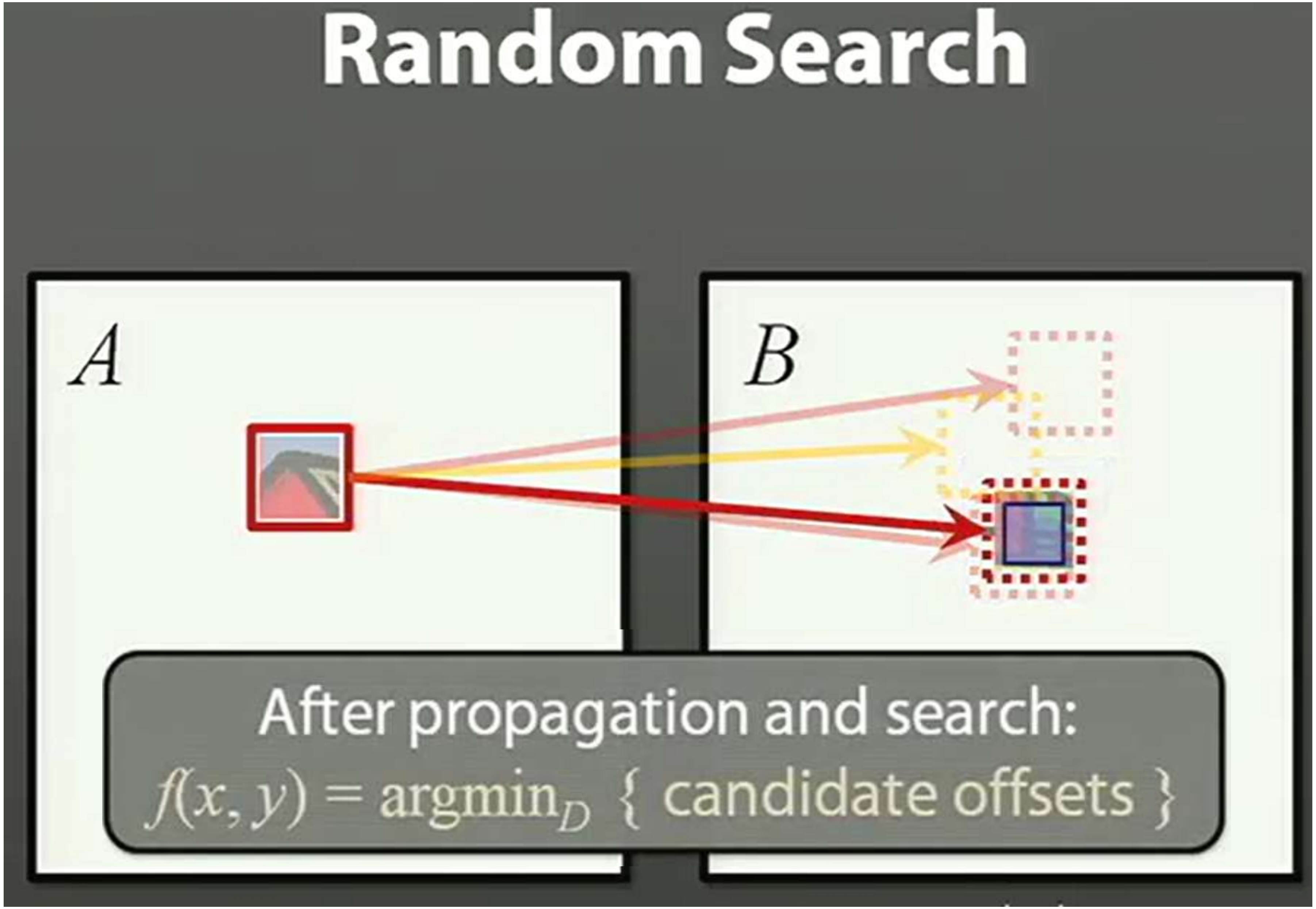
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



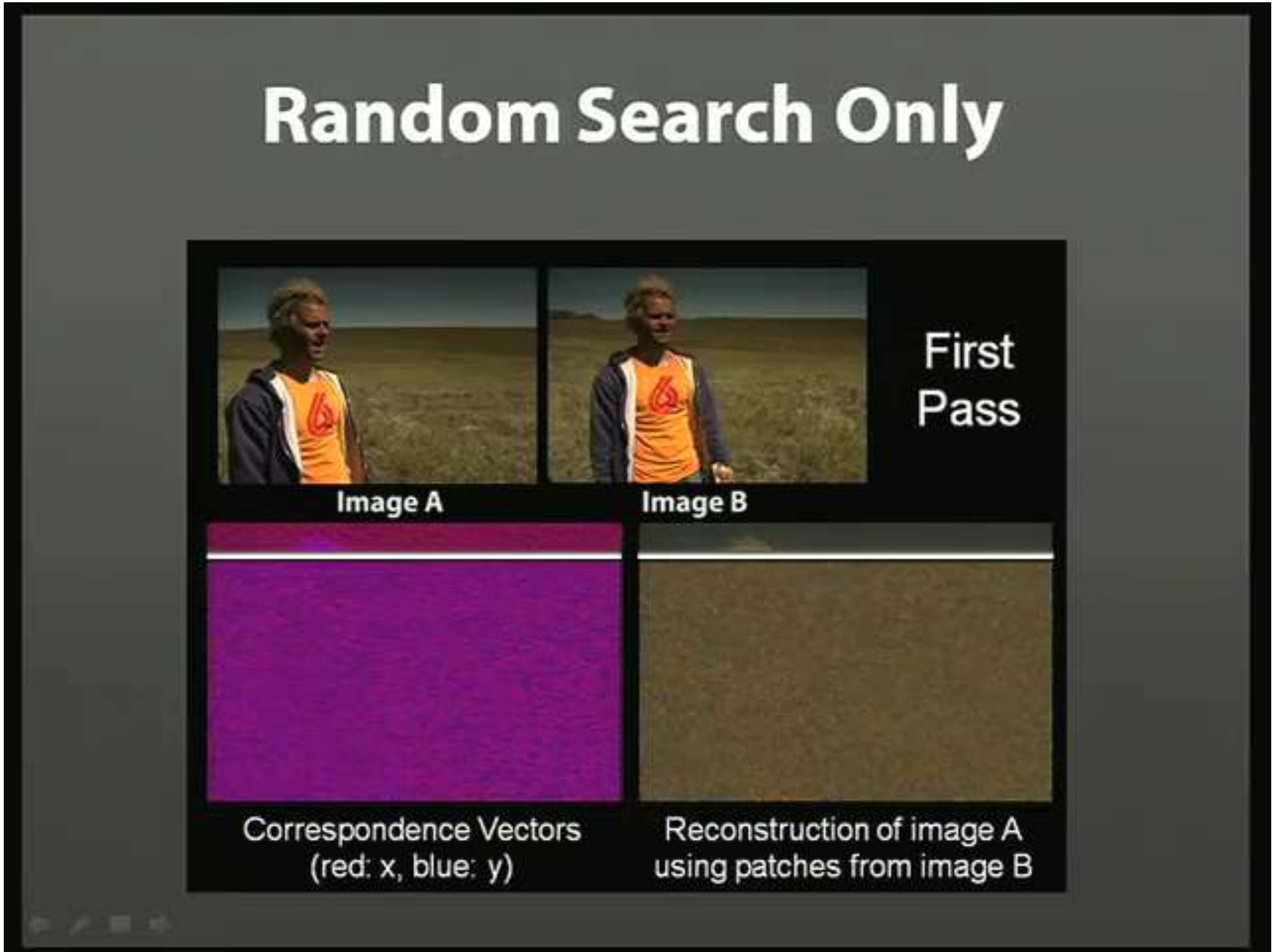
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



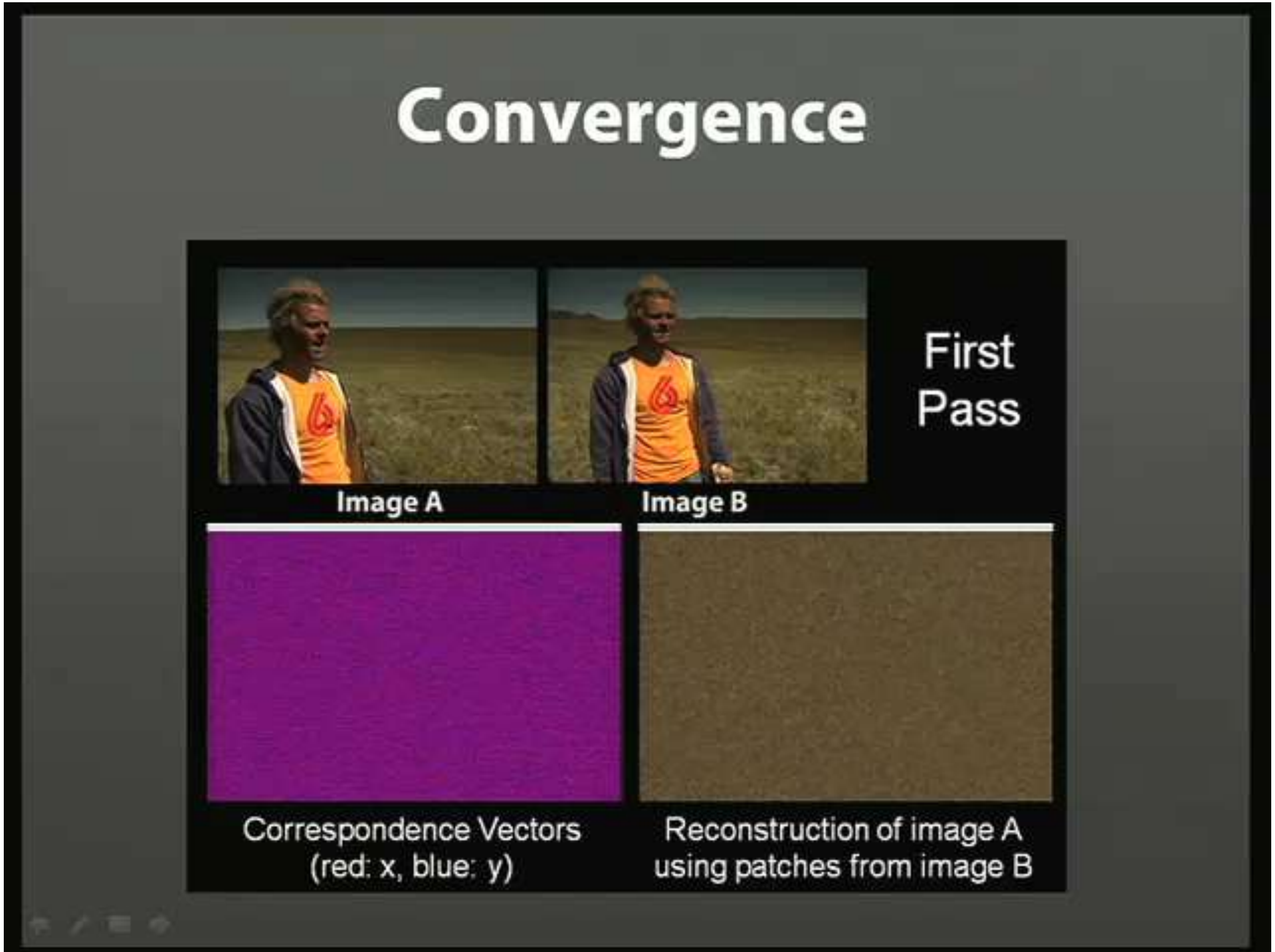
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



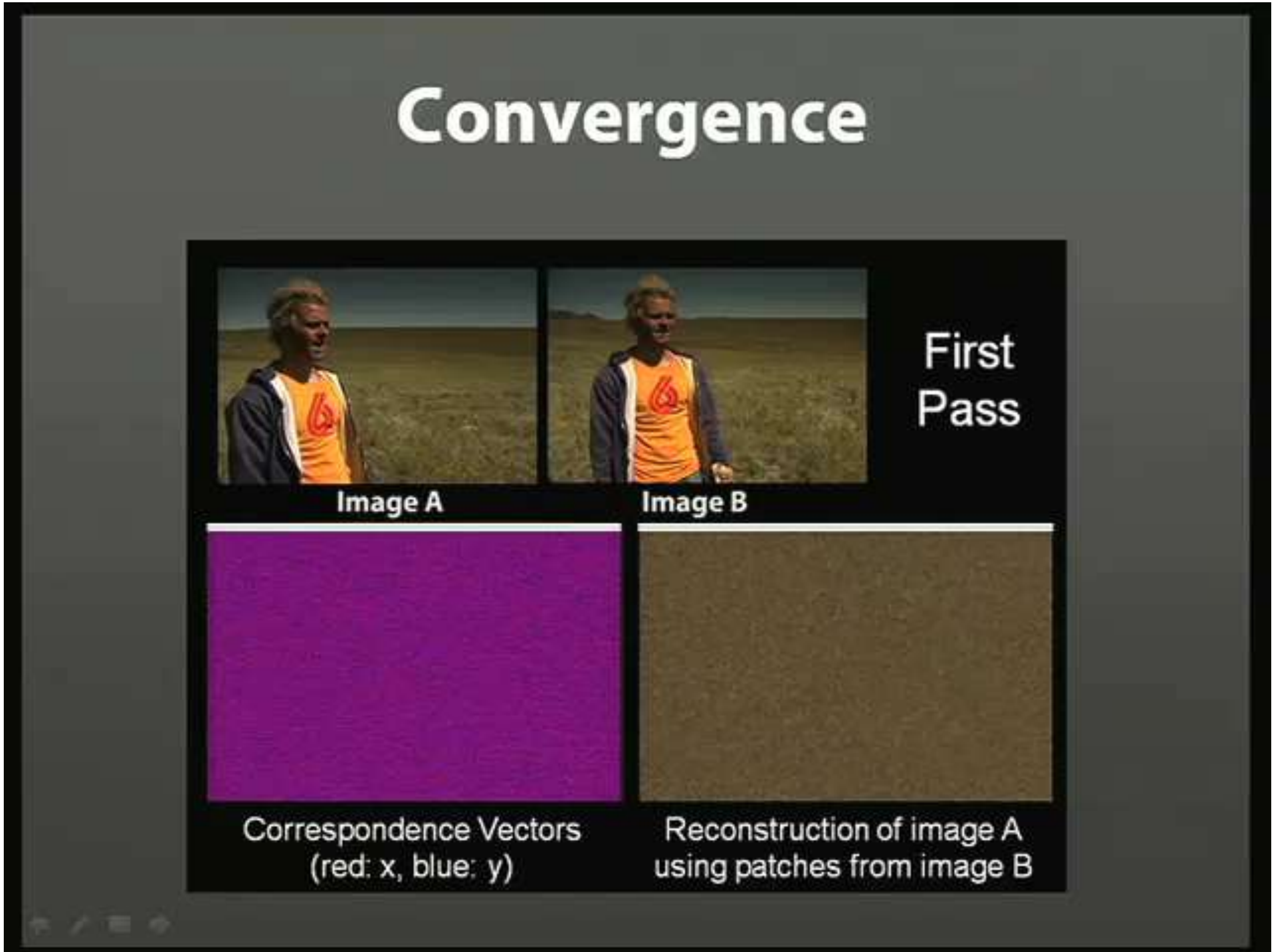
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.

小作业安排

- 具体代码和要求会放在canvas上
- VAE
  - 第一周周日放出
  - 第四周周日截止
- GAN
  - 第二周周日放出
  - 第六周周日截止
- Diffusion Model
  - 第三周周日放出
  - 第八周周日截止

大作业——12周截止

- 不会提供公共计算资源

- 分组，最多每组5人
  - 分组在**第四周**确定，之后不得更改

- 课题：微调文生图/视频模型/定制化生成
  - 推荐使用DreamBooth/Textual Inversion/Imagic/Custom Diffusion/Wan/AnimateDiff
  - 每组内部用统一模型，提交统一报告
  - 每个组员使用自己图片/视频，组员之间不可以重复，每个同学报告里提供自己生成的图像/视频

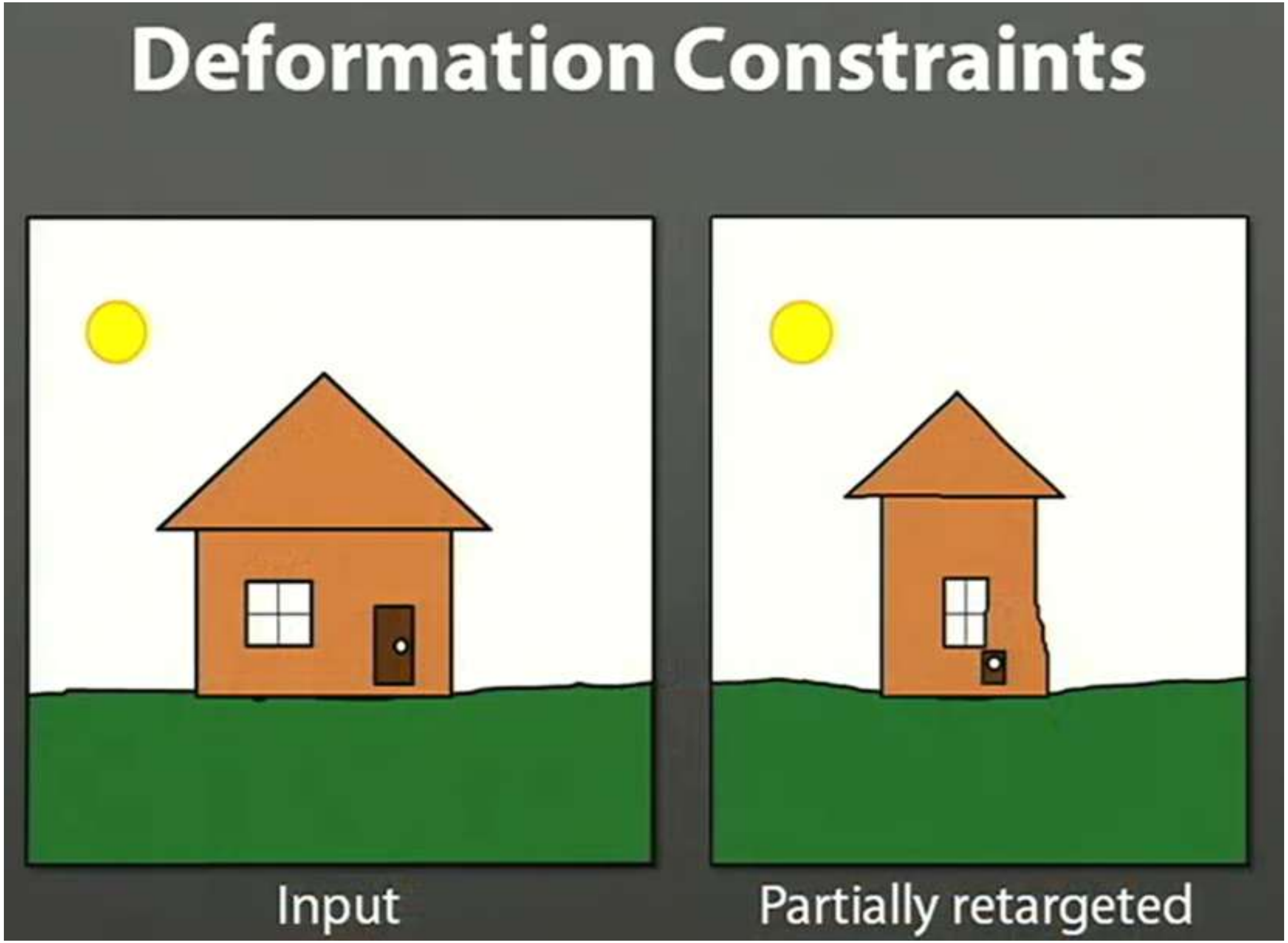
- 可以自行定义与课程相关的题目
  - 如自行选题，需要在**第四周**前联系确定是否合适

PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.

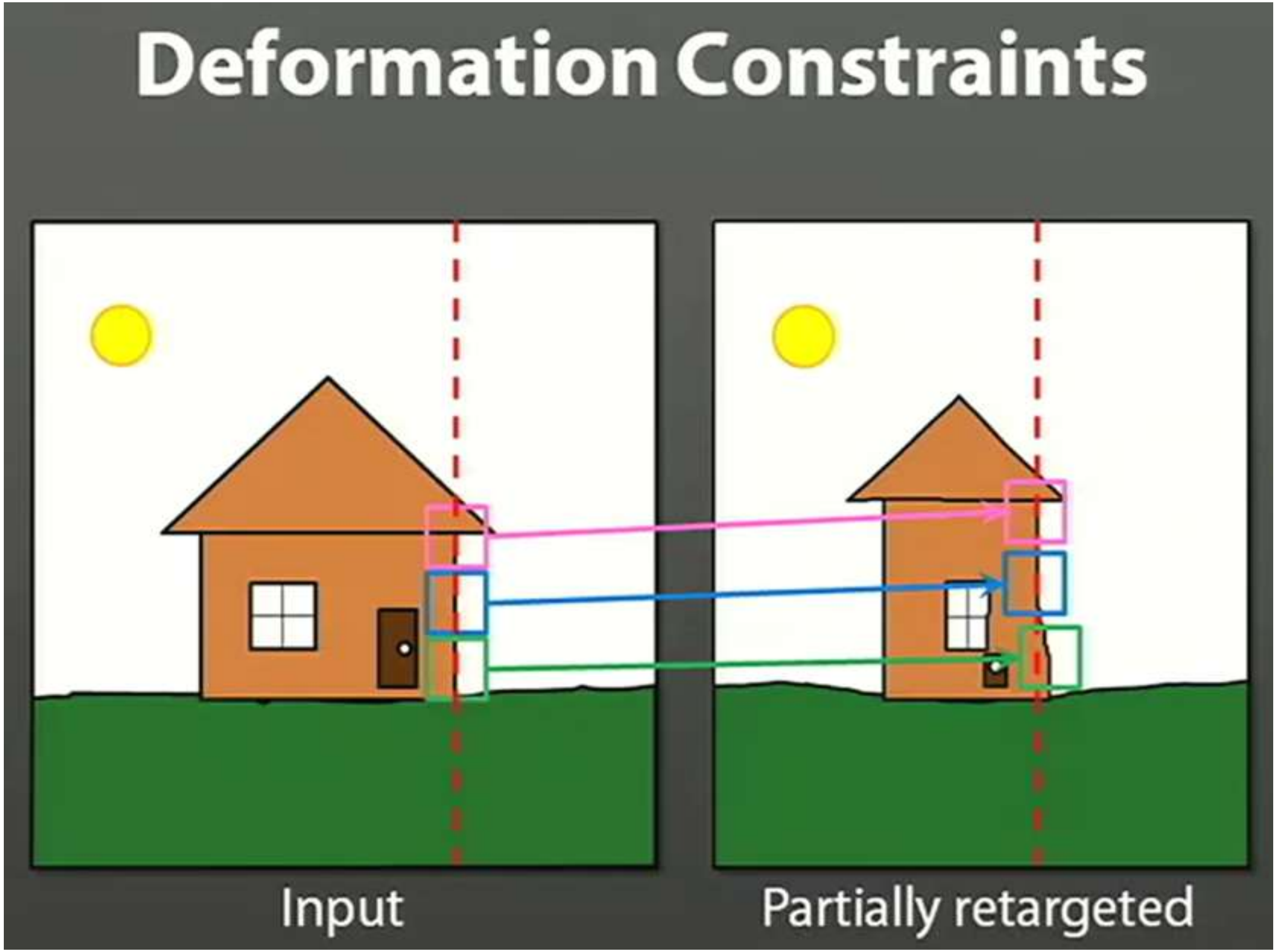
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



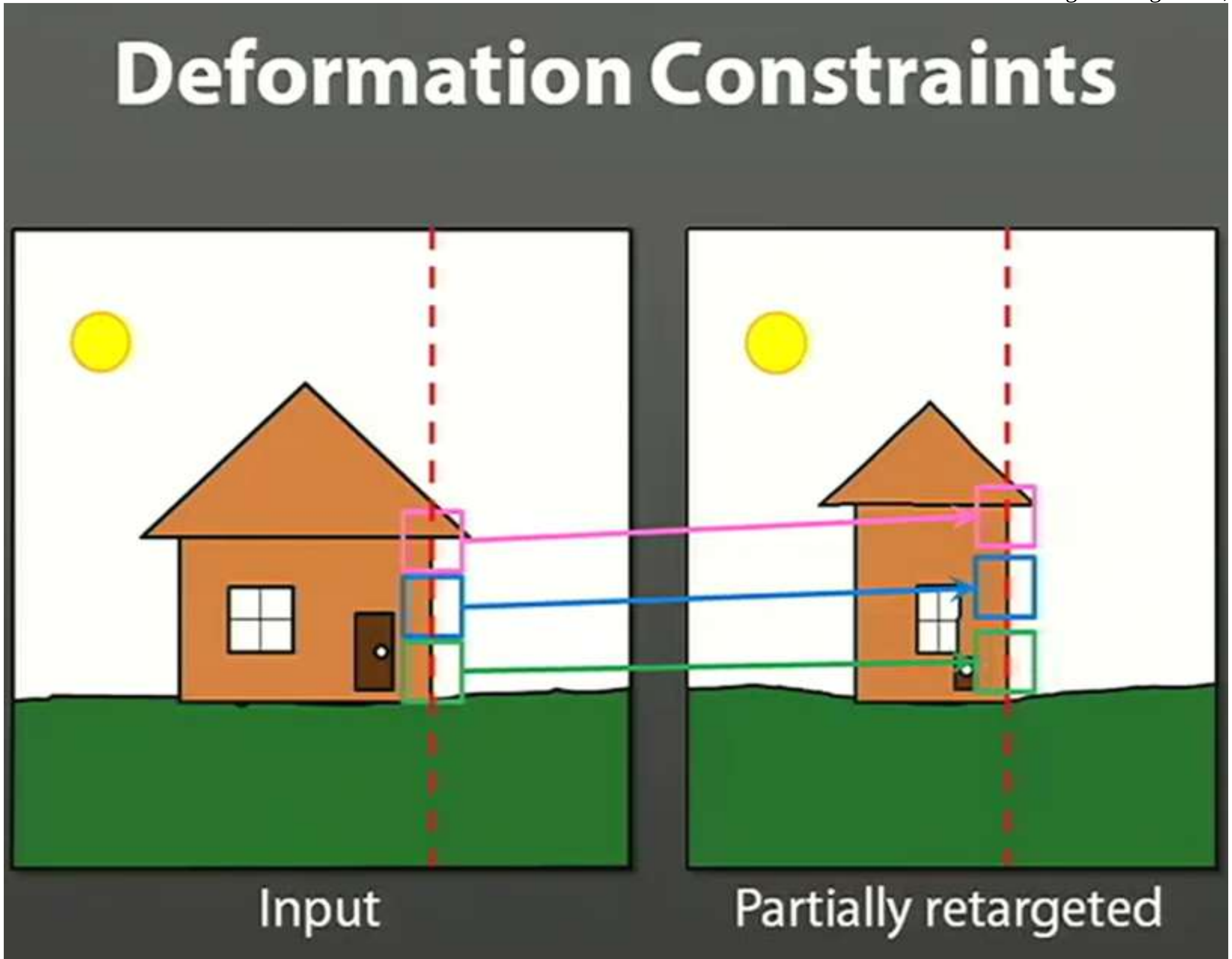
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



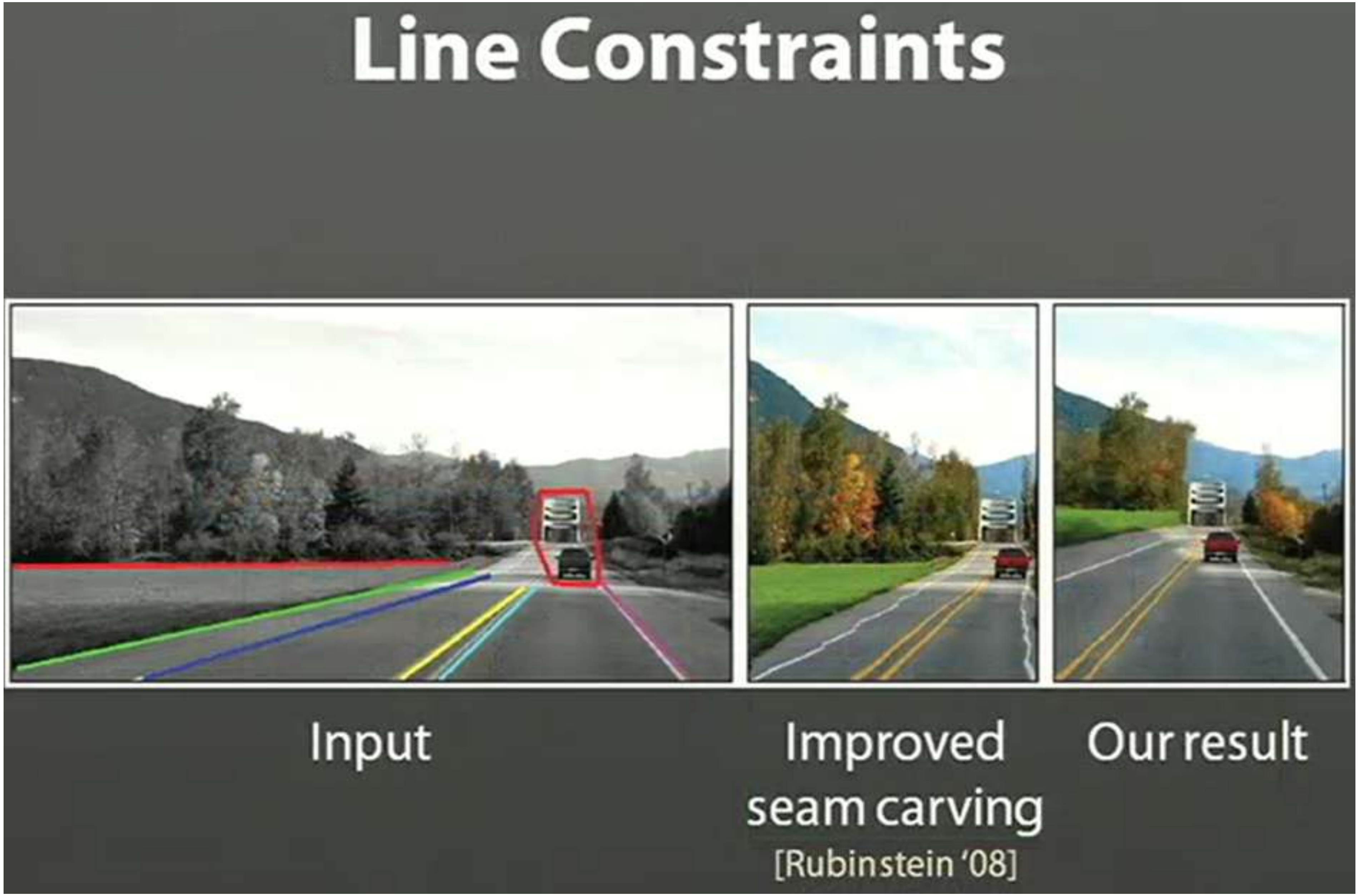
PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.

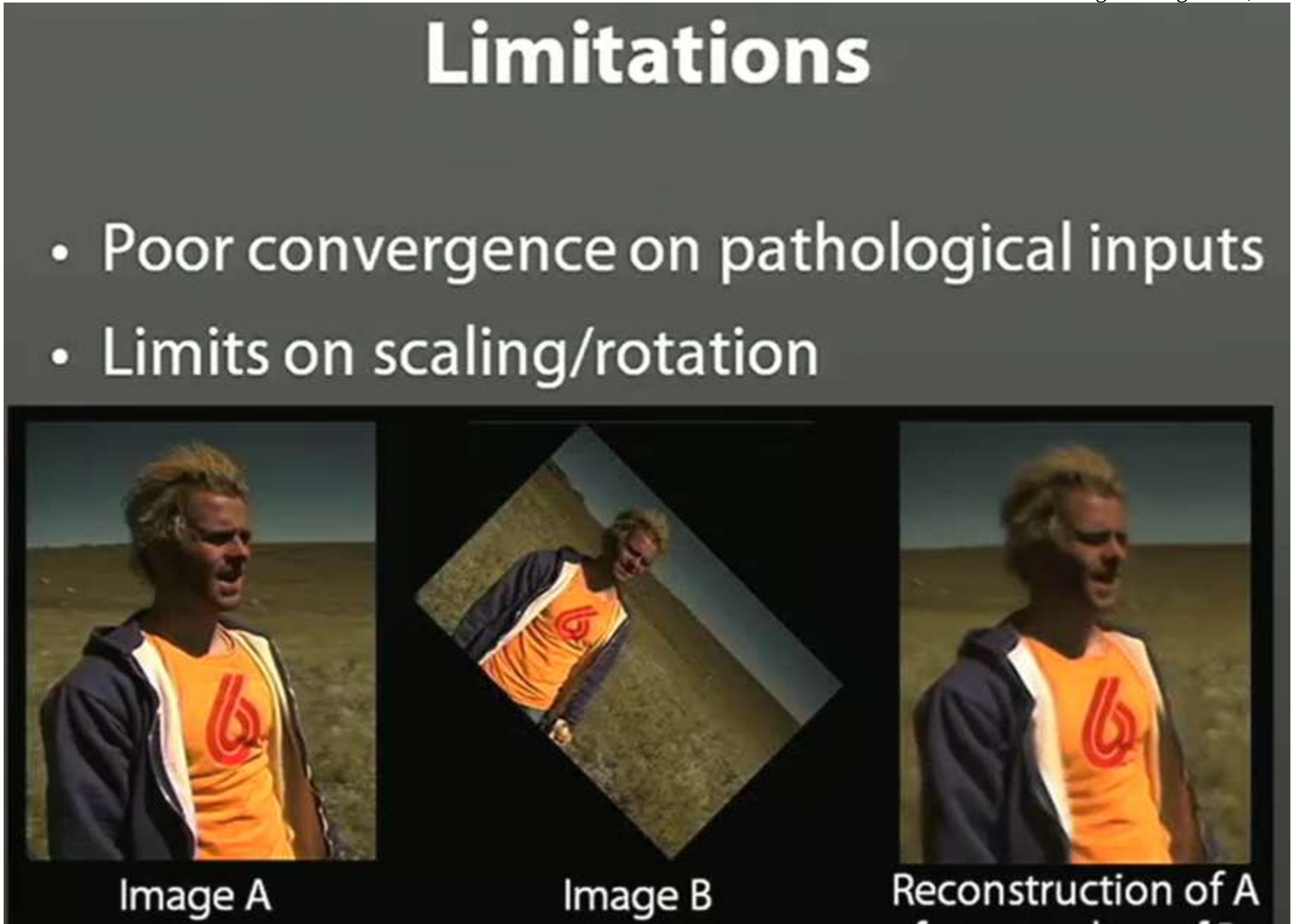


PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.

PatchMatch

* C. Barnes, E. Shechtman, A. Finkelstein, and D. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. TOG, 2009.



PatchMatch

- SoTA image editing 10 years ago

- We can see some metaphors…
  - Propagation: CNN, Autoregressive Models
  - Random Search：Diffusion Models, Transformer
  - ControlNet

- But there are no new content generated…
  - Just algorithms, not models (without learning)

Deep Image Prior

Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

In image restoration problems the goal is to recover original image x having a corrupted image x_0. Such problems are often formulated as an optimization task:

$$\min_x E(x; x_0) + R(x)\,, \tag{1}$$

where $E(x; x_0)$ is a *data term* and $R(x)$ is an *image prior*. The data term $E(x; x_0)$ is usually easy to design for a wide range of problems, such as super-resolution, denoising, inpainting, while image prior $R(x)$ is a challenging one. Today's trend is to capture the prior $R(x)$ with a ConvNet by training it using large number of examples.

We first notice, that for a surjective $g : \theta \mapsto x$ the following procedure in theory is equivalent to (1):

$$\min_\theta E(g(\theta); x_0) + R(g(\theta))\,.$$

In practice g dramatically changes how the image space is searched by an optimization method. Furthermore, by selecting a "good" (possibly injective) mapping g, we could get rid of the prior term. We define $g(\theta)$ as $f_\theta(z)$, where f is a deep ConvNet with parameters θ and z is a fixed input, leading to the formulation

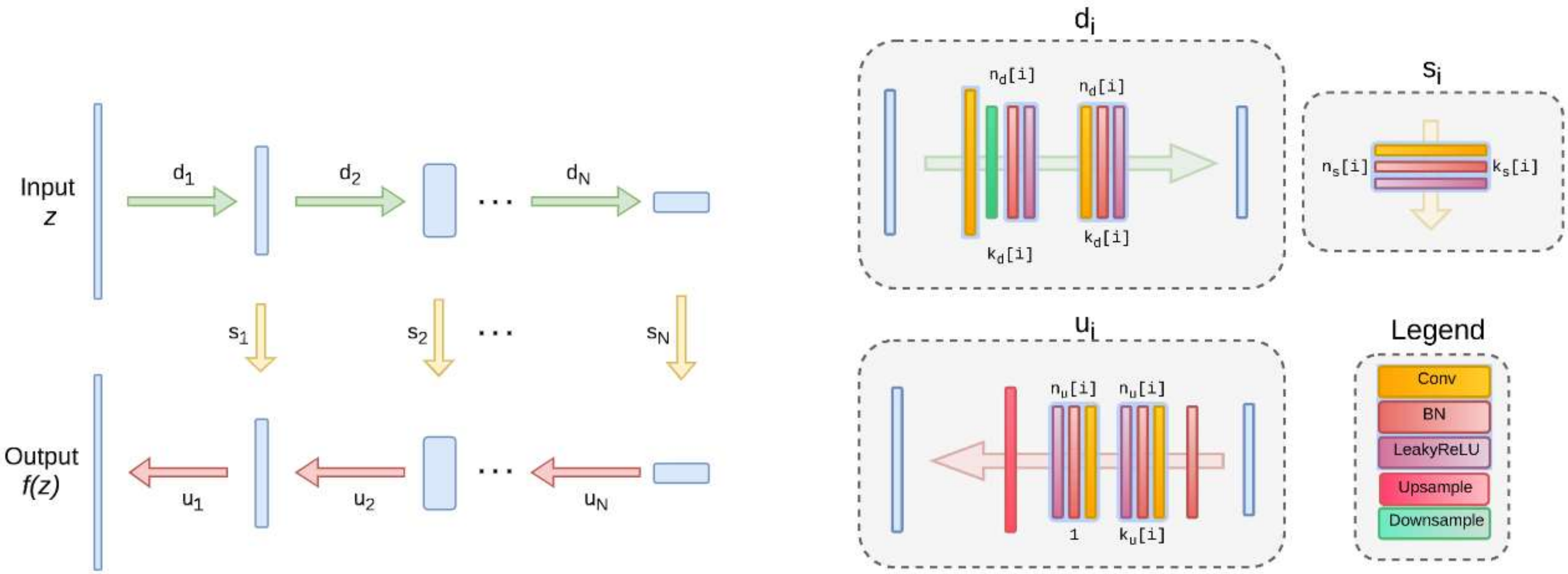
$$\min_\theta E(f_\theta(z); x_0)\,.$$

Here, the network f_θ is initialized randomly and input z is filled with noise and fixed.

In other words, **instead of searching for the answer in the image space we now search for it in the space of neural network's parameters**. We emphasize that we never use a pretrained network or an image database. Only corrupted image x_0 is used in the restoration process.

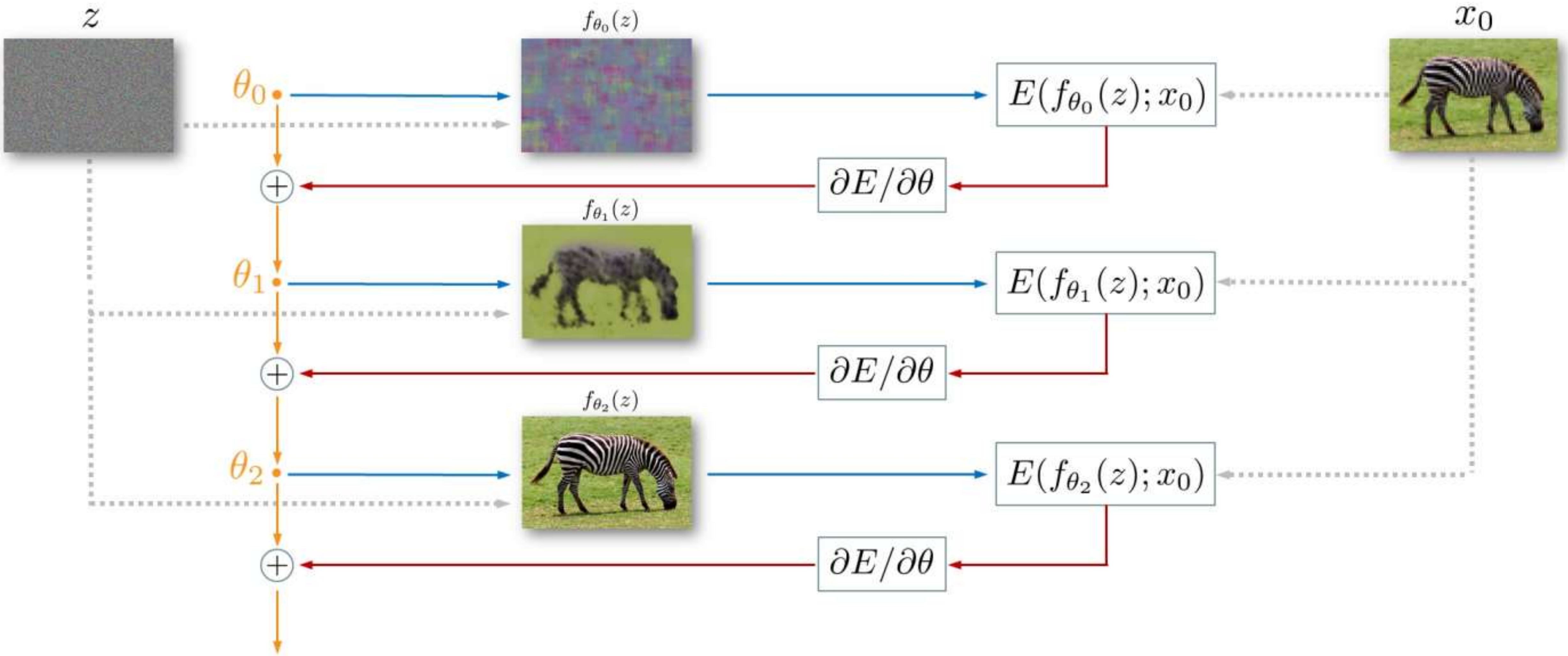
Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.



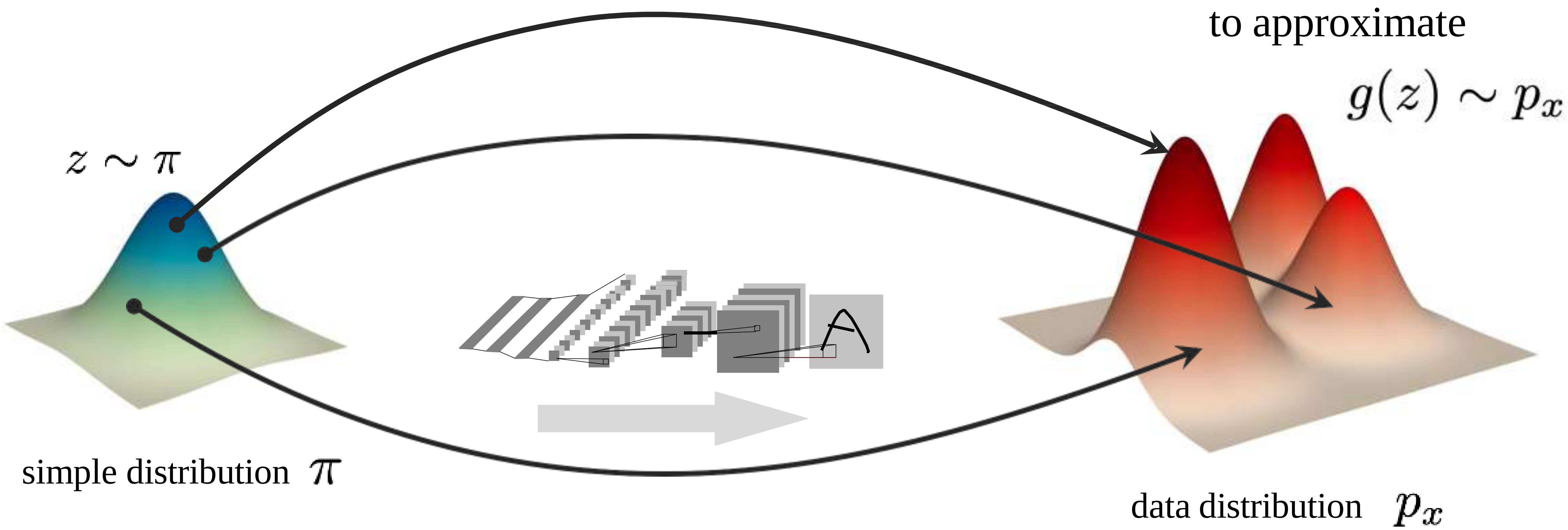
Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.



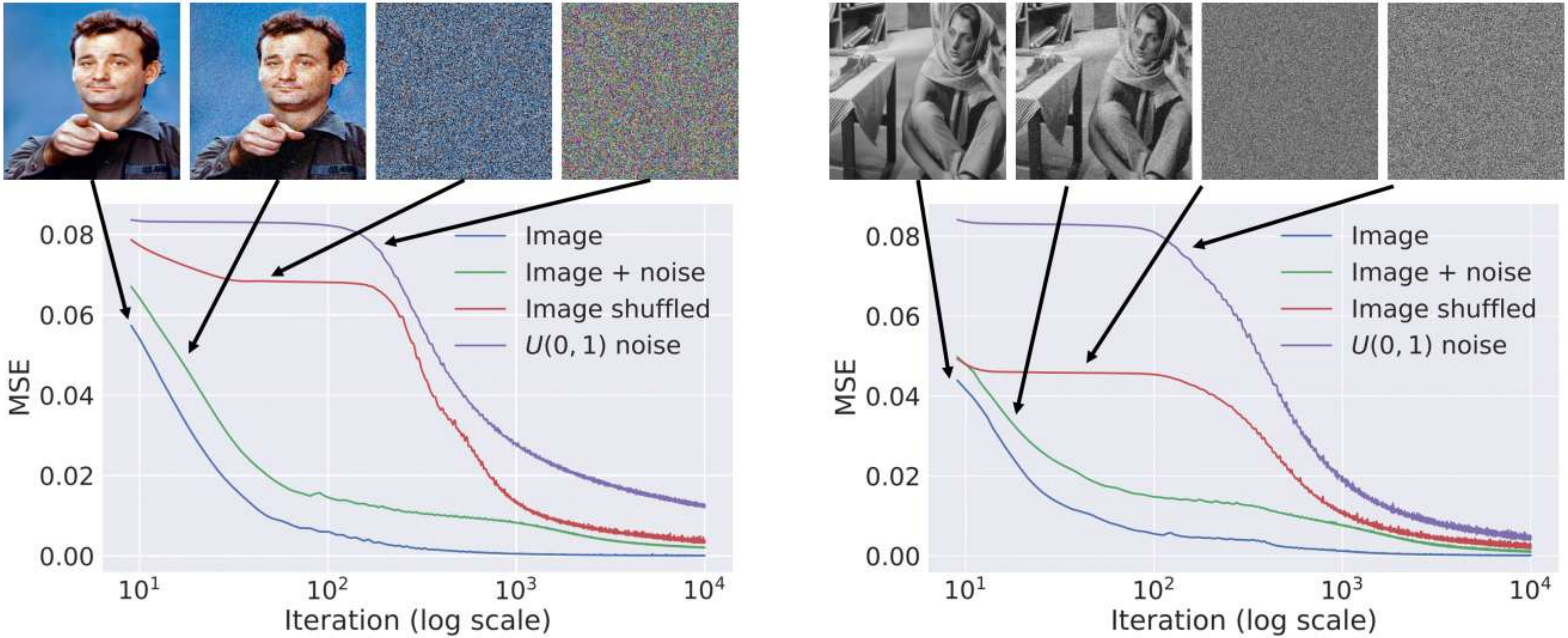
Recap.

- Deep image prior uses individual network for each image

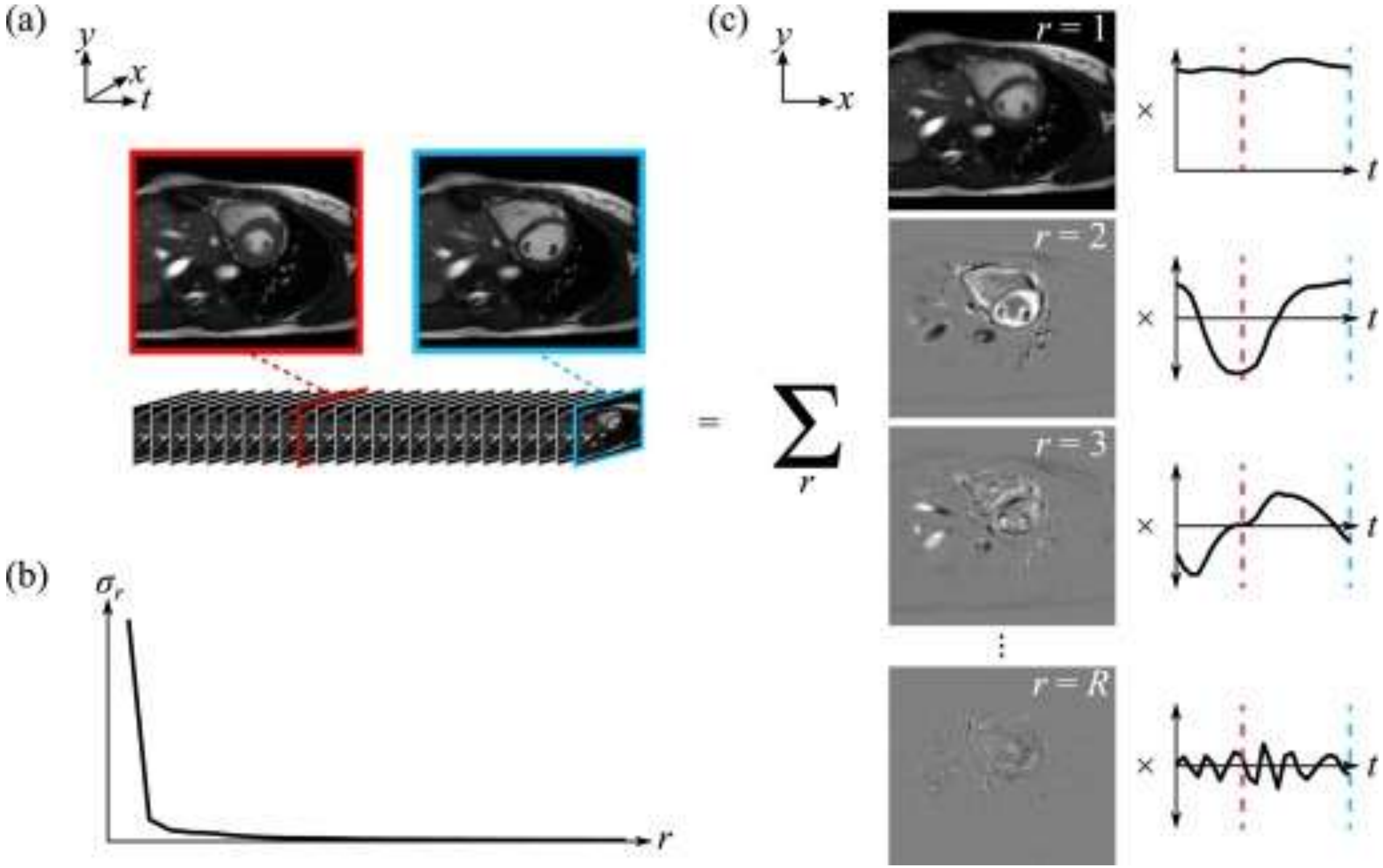


Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.



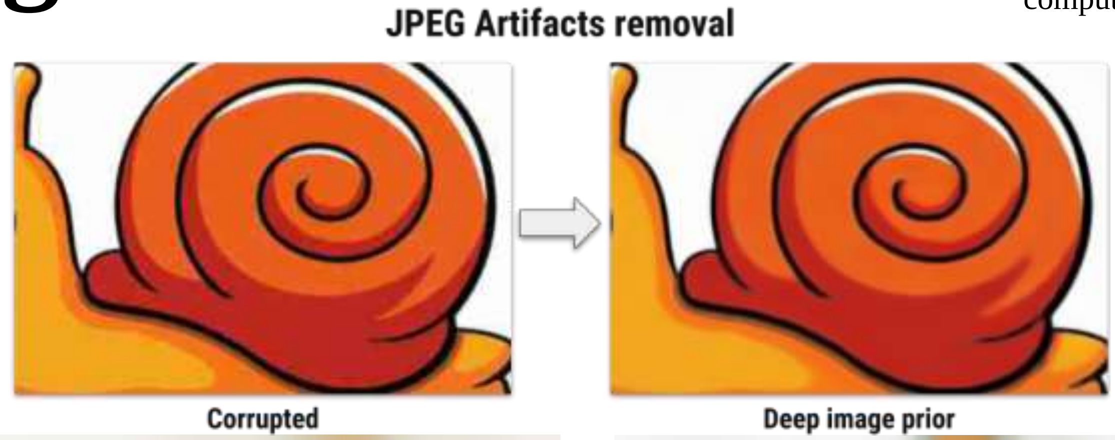
Compressed Sensing / Low-Rank Approximation

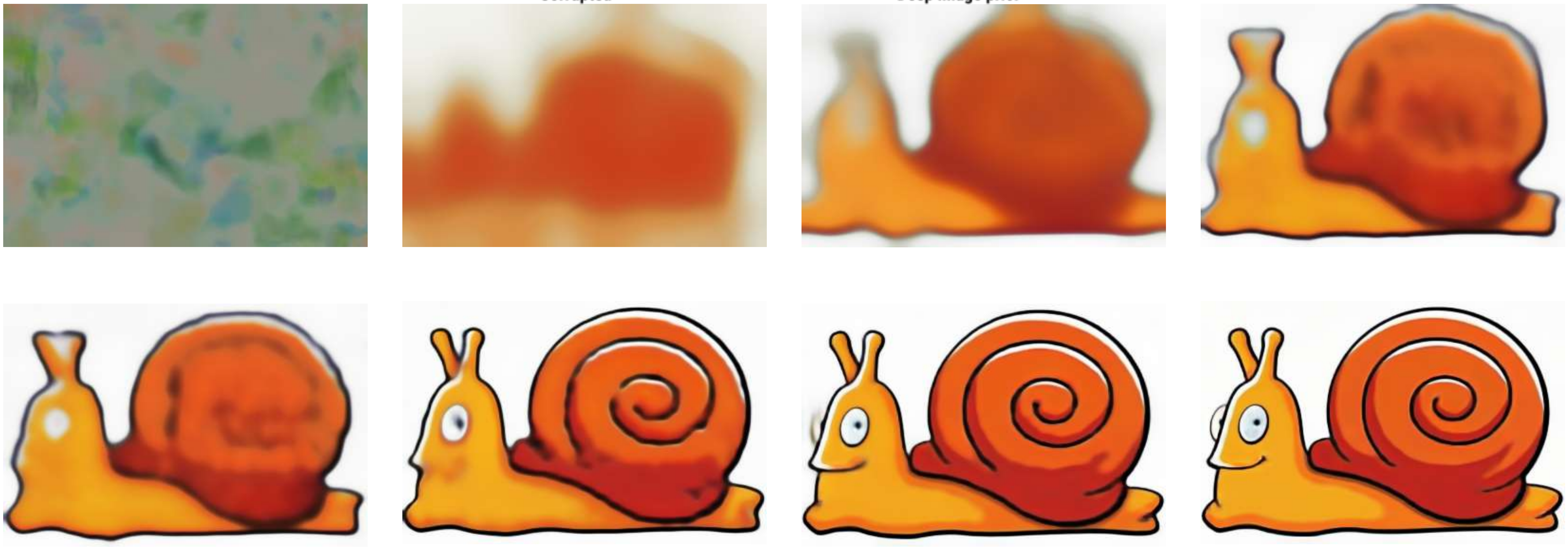


https://en.wikipedia.org/wiki/Low-rank_approximation

Deep Image Prior

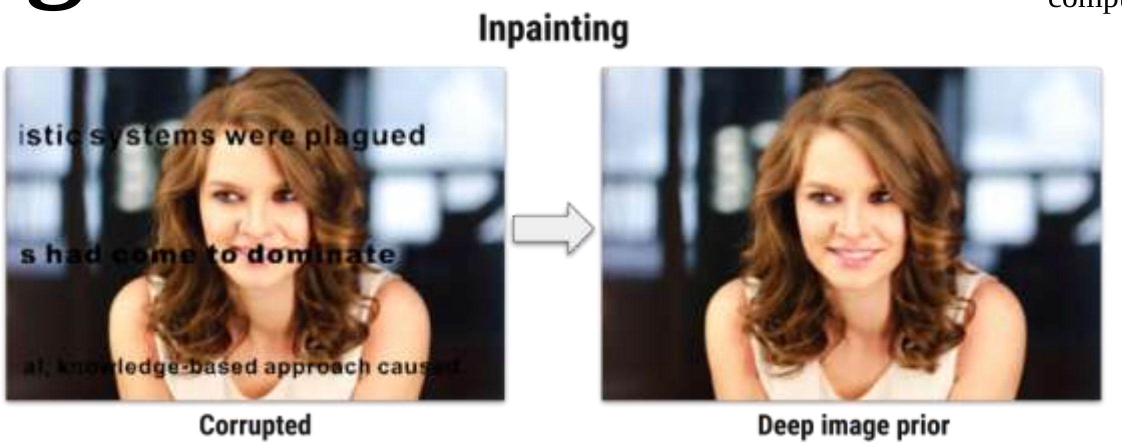
Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

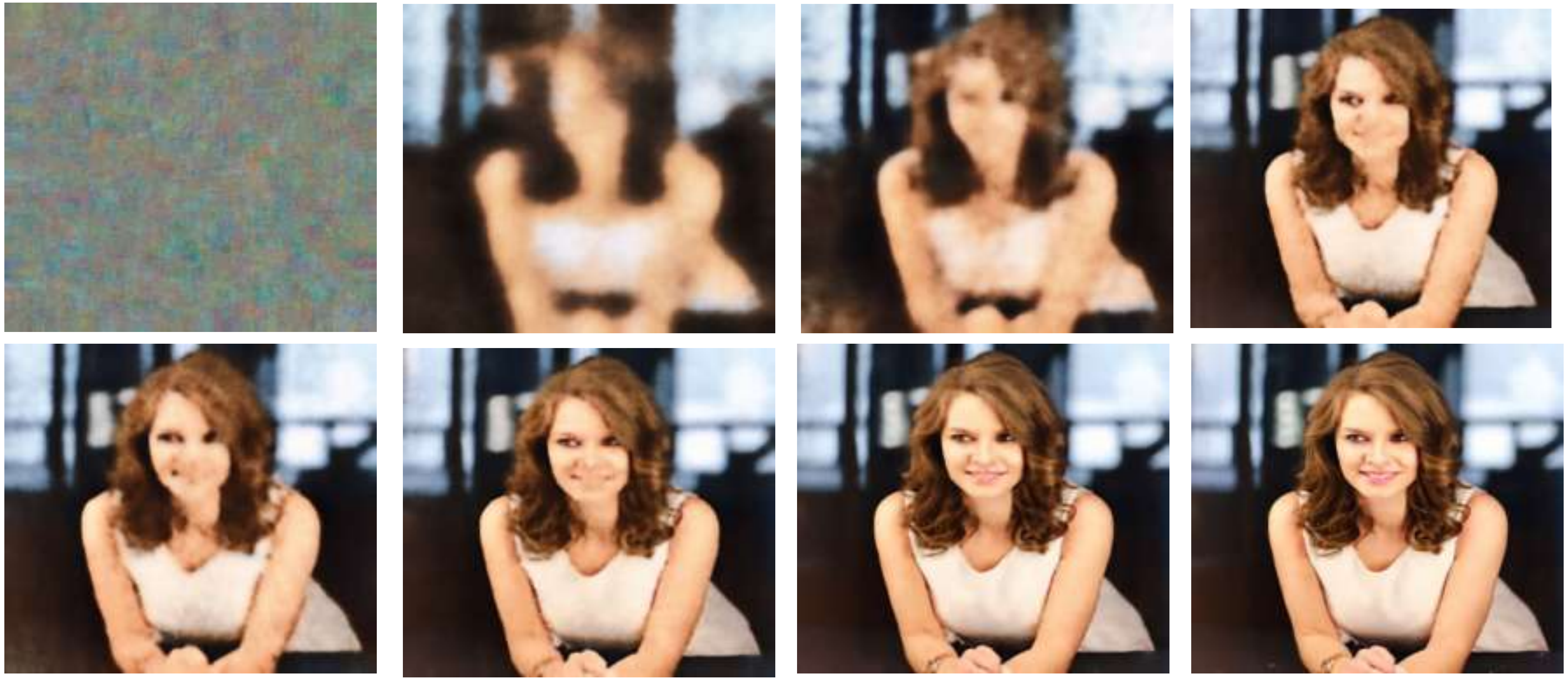




Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.





Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

(a) Original image

(b) Bicubic, **Not trained**

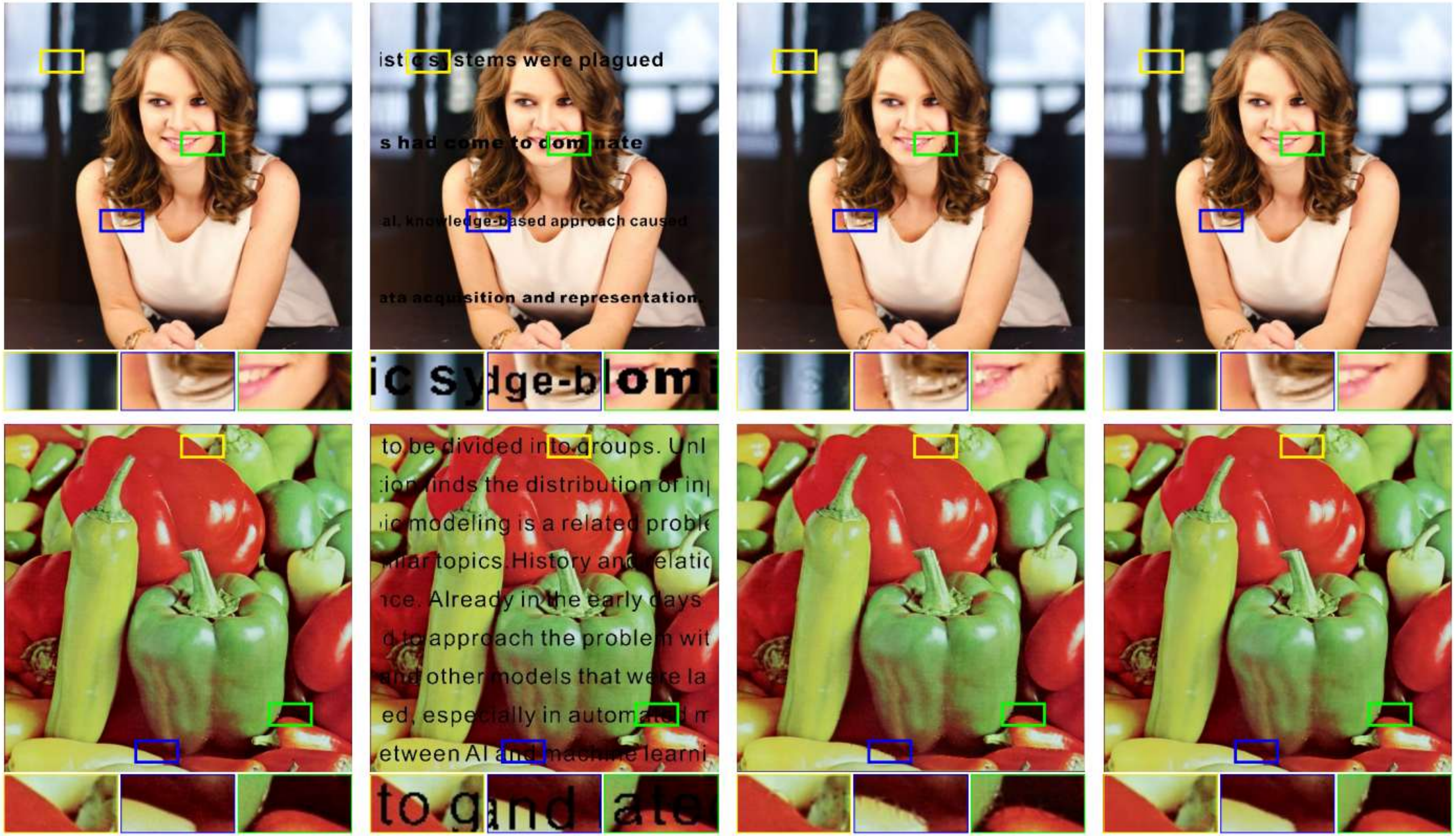
(c) Ours, **Not trained**

(d) LapSRN, **Trained**

(e) SRResNet, **Trained**

Deep Image Prior

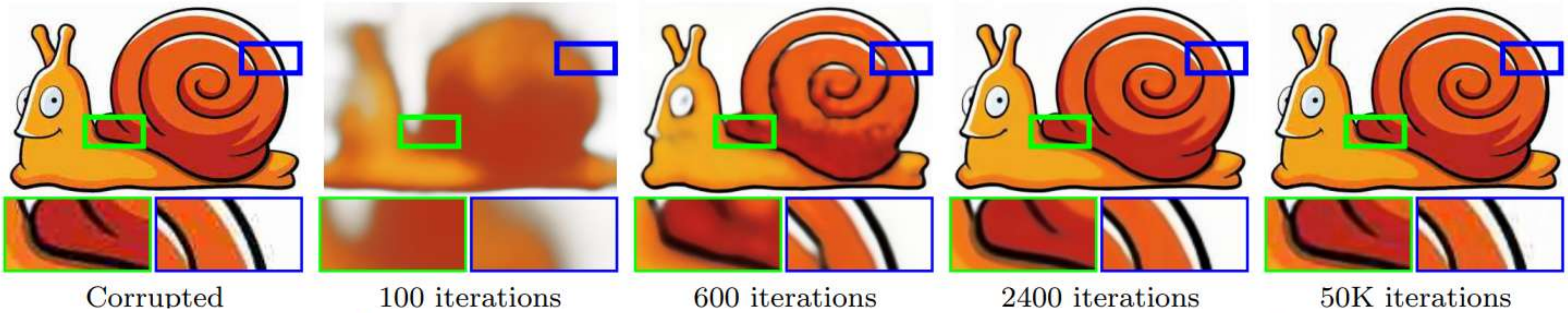
Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

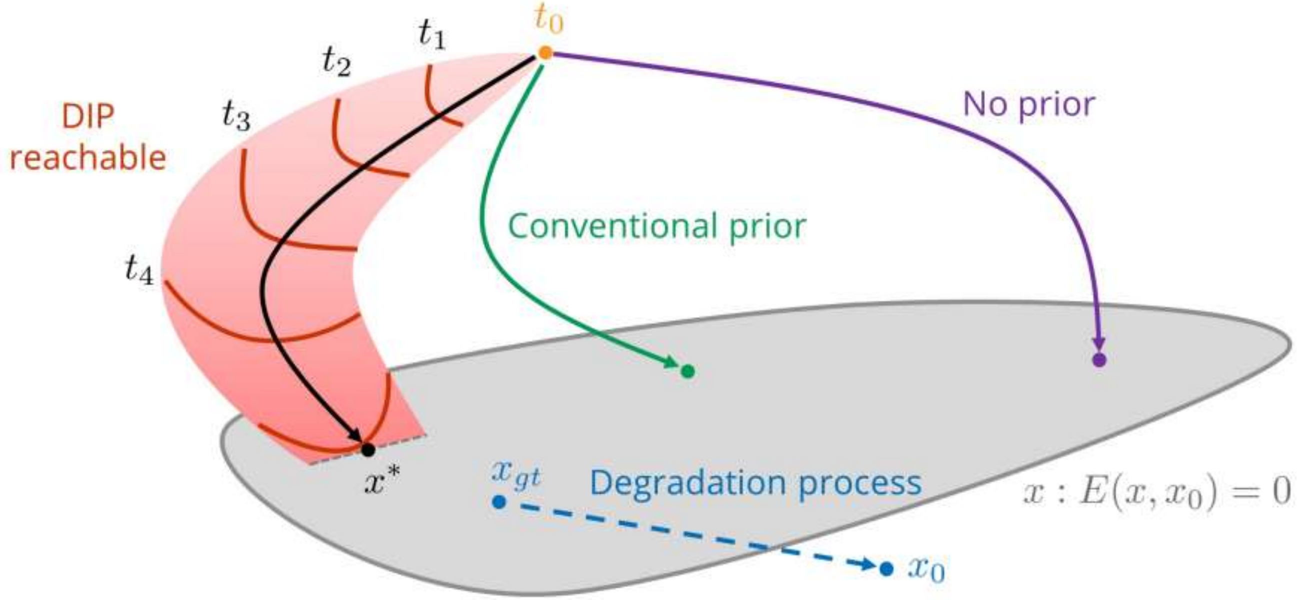


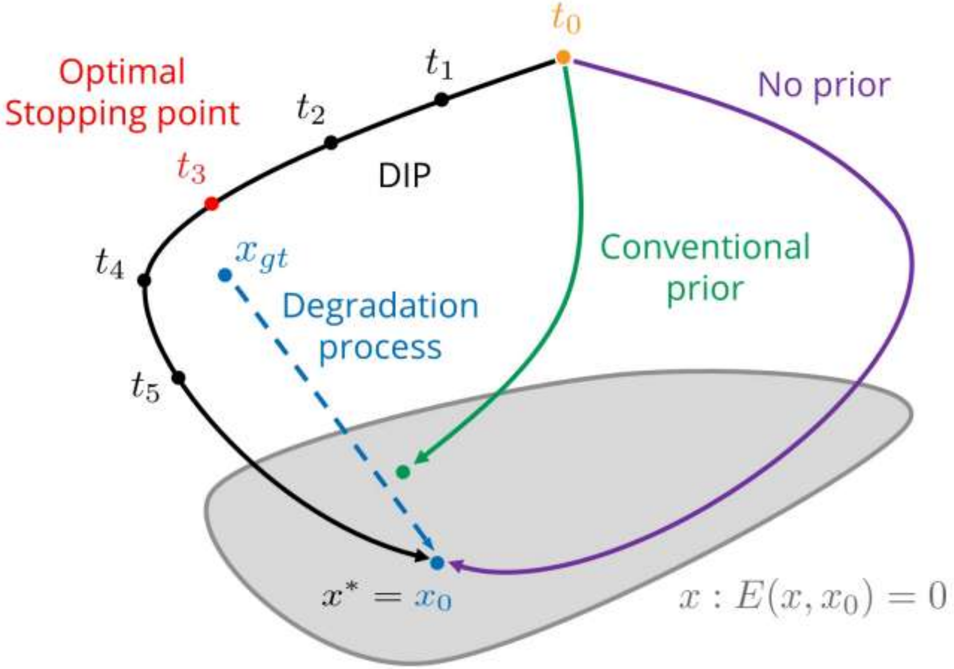
(a) Original image (b) Corrupted image (c) Shepard networks [44] (d) Deep Image Prior

Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

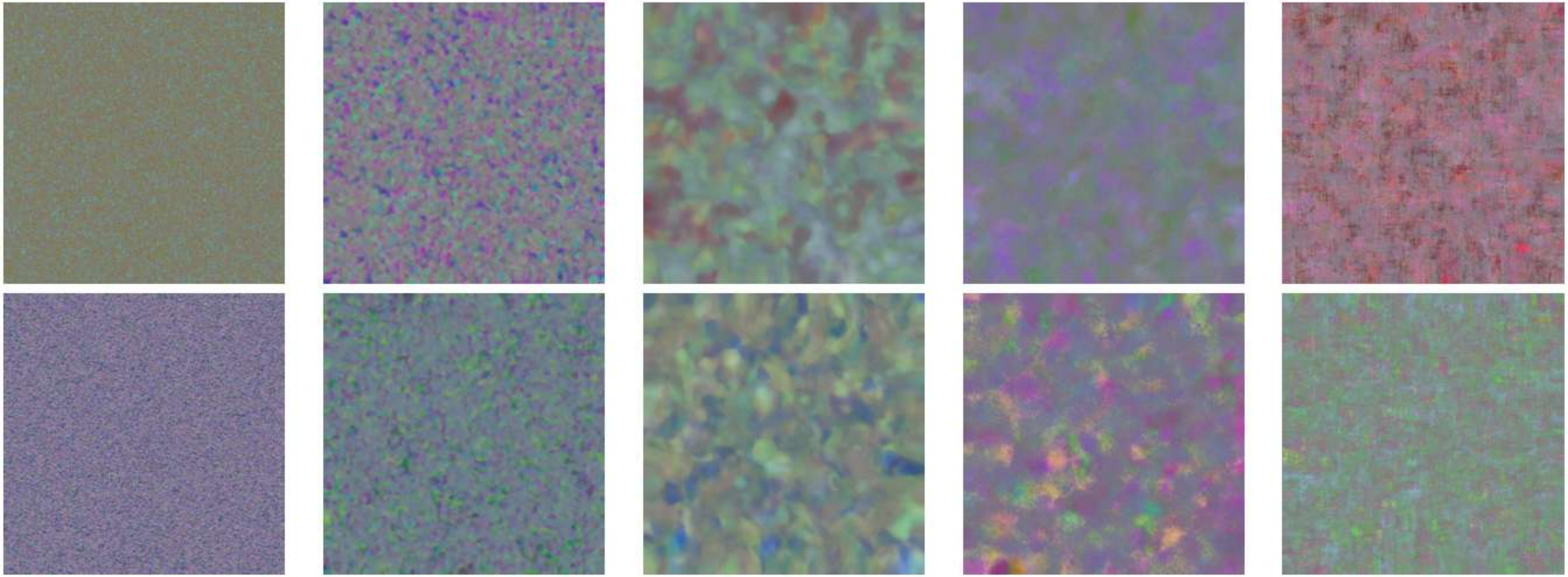






Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.



a) Hourglass-1 b) Hourglass-3 c) Hourglass-5 d) Skip-5 e) Skip-5-nearest

Deep Image Prior

Ulyanov, Dmitry, Andrea Vedaldi, and Victor Lempitsky. "Deep image prior." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

(a) Input (white=masked)

(b) Encoder-decoder, depth=6

(c) Encoder-decoder, depth=4

(d) Encoder-decoder, depth=2

(e) ResNet, depth=8

(f) U-net, depth=5

Deep Image Prior

- Simplest/naïve generator, now we have models!
- Do not depend on large scale datasets
- Deep networks have their priors of images

- Slow…
- Still no new content generated

- Related works
  - Tensor Decomposition / Compressed Sensing
  - SinGAN (ICCV 2019 best paper)