Video Inpainting of Complex Scenes

Andrés Almansa

Alasdair Newson – Yann Gousseau – Matthieu Fradet – Patrick Perez

Image and video inpainting

What is inpainting?

- Removal and filling of a region in an image or video
- The inpainted region should be visually convincing/pleasing

Image to inpaint

Inpainted image

What is inpainting useful for?

- Restoring/improving/modifying images/videos
- Post-production of films

Original

Restored

Video Inpainting

Inpainted video

Video Inpainting

Original video

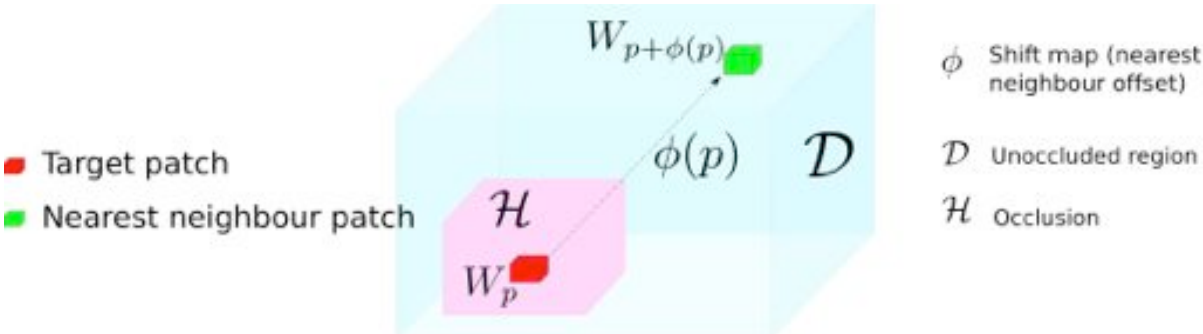
Video Inpainting

What are the challenges of video inpainting?

- Moving object reconstruction
- Simultaneous foreground/background reconstruction
- Inpainting with moving background
- Extremely long computational times

Inpainting example (from Wexler *et al.* 2007)

Video Inpainting



W_p: a patch centered at p

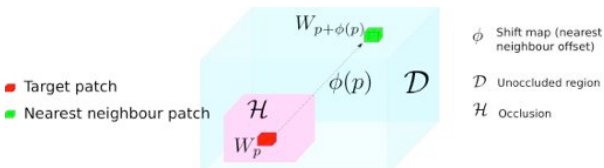
Video Inpainting

Inpainting Principle

Input: $u|_{\mathcal{D}}$ Output: $u|_{\mathcal{H}}$

Find $u|_{\mathcal{H}}$ by minimizing

$$E(u,\phi) = \sum_{p\in\mathcal{H}} ||W_p^u - W_{p+\phi(p)}^u||_2^2$$



W_p: a patch centered at p

Challenges

- non-convex energy
- high dimensionality (dimension $= 5 \times 5 \times 5 \times 3 \approx 500$)

Solutions

- alternate (convex) minimization w.r.t. u and ϕ
- coarse-to-fine processing
- approximate nearest neighbours
- fine-level texture features in coarsest level

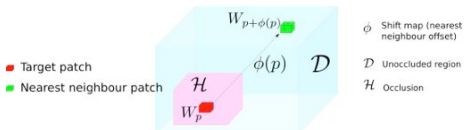
Video Inpainting – *Alternate Minimization*

Inpainting Principle

Input: $u|_{\mathcal{D}}$ Output: $u|_{\mathcal{H}}$

Find $u|_{\mathcal{H}}$ by minimizing

$$E(u,\phi) = \sum_{p \in \mathcal{H}} ||W_p^u - W_{p+\phi(p)}^u||_2^2$$



W_p: a patch centered at p

Algorithm (inspired by † and ‡):

Alternate Minimization on u and ϕ:

$u^0 \leftarrow$ Initialisation$(u|_{\mathcal{D}}, \mathcal{H})$

1/ $\phi^{k+1} \leftarrow$ NearestNeighbourSearch(u^k)

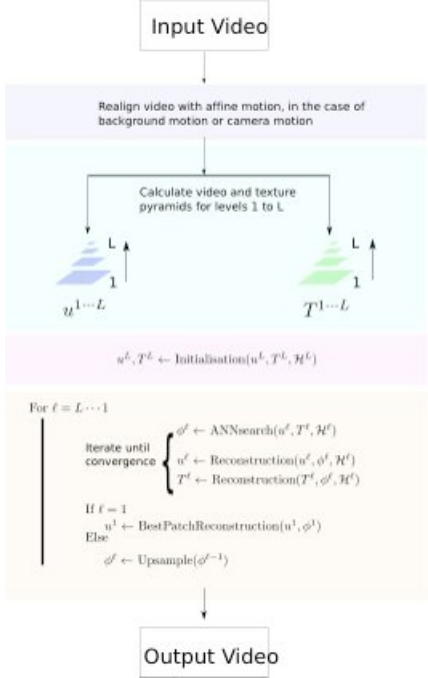
2/ $u^{k+1} \leftarrow$ VideoReconstruction(ϕ^{k+1})

(aggregation of patches)

† *Y. Wexler, E. Schechtman, M. Irani*, **Space-Time Completion of Video**, *PAMI 2007*

‡ *P. Arias, G. Facciolo, V. Caselles, G. Sapiro*, **A Variational Framework for Exemplar-Based Image Inpainting**, *IJCV 2011*

Video Inpainting
- *Coarse to Fine*



Video Inpainting – *Approximate Nearest Neighbor Search*

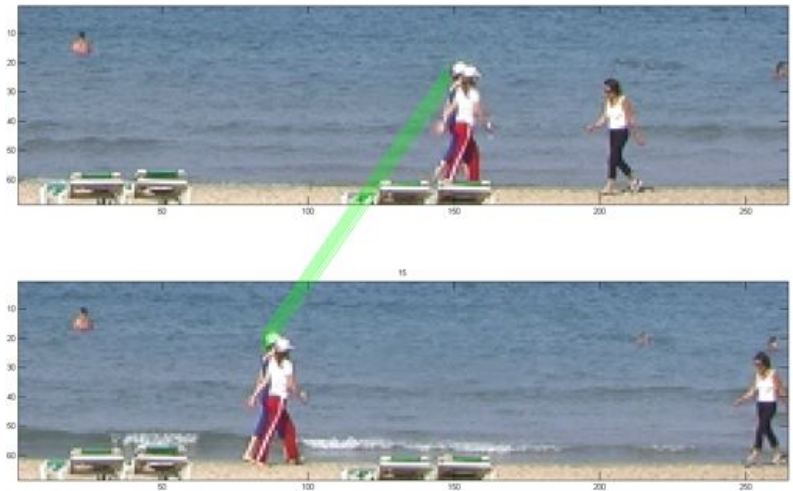
High dimensionality of problem means NN search is *very* slow.

- Previously used ANN search algorithm (kdTrees) very slow
- We extend the PatchMatch (Barnes *et al.* 2009) algorithm to spatio-temporal case.
- PatchMatch: ANN search algorithm for image patches

† *C. Barnes, E. Schechtman, A. Finkelstein, D. B. Goldman,* **PatchMatch: a randomized correspondence algorithm for structural image editing**, *ACM Transactions on Graphics (2009)*

Video Inpainting – *Approximate Nearest Neighbor Search*

PatchMatch based on piecewise constancy of the shift map

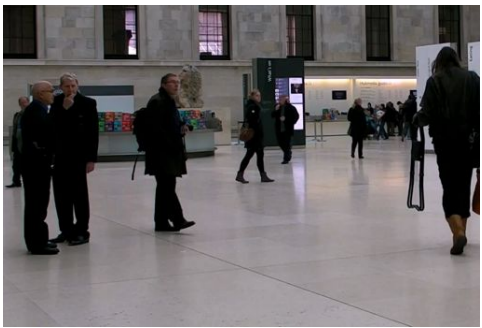


Visual comparisons
Granados et al.

- 10-50 times speedup with 3D PatchMatch
- 10 times speedup compared to Granados *et al.*

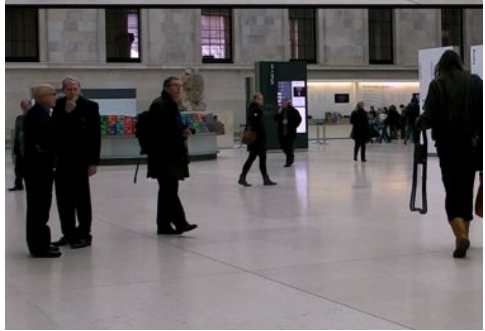
M. Granados, J. Tompkin, K.I. Kim, O. Grau, J. Kautz, C. Theobalt,
How Not to Be Seen - Object Removal from Videos of Crowded Scenes,
Computer Graphics Forum, 2012

Original with occlusion



Granados

ours



Dealing with textures in images and videos

Why do textures pose a problem ?

Original image

Dealing with textures in images and videos

Why do textures pose a problem ?

Inpainted image

Dealing with textures in images and videos

Why do textures pose a problem ?

Incorrect approximate nearest neighbours

Modified patch distance

We wish to include some information pertaining to the texture.

Idea : include an estimation of the local variance

Different possibilities were tested. Finally, we chose (inspired by Liu and Caselles 2013[†]) :

$$\text{SSD: } [\mathrm{R}, \mathrm{G}, \mathrm{B}, \alpha|\nabla_x I|_\nu, \alpha|\nabla_y I|_\nu]$$

α: a weighting scalar

$|\nabla_\bullet I|_\nu$: average absolute value of the $\bullet$ - derivative in a neighbourhood ν of size $|\nu|$.

[†] *Y. Liu, V. Caselles*, **Exemplar-Based Image Inpainting Using Multiscale Graph Cuts**, *IEEE TIP (2013)*

Modified patch distance

Example of image created by $|\nabla_x I|_\nu$

Modified patch distance

Example of the impact of the modified distance

PatchMatch with regular SSD

Modified patch distance

Example of the impact of the modified distance

PatchMatch with modified SSD

Image example

Inpainting with unmodified patch distance

Image example

Inpainting with "Image Melding" (Darabi *et al.* 2012)

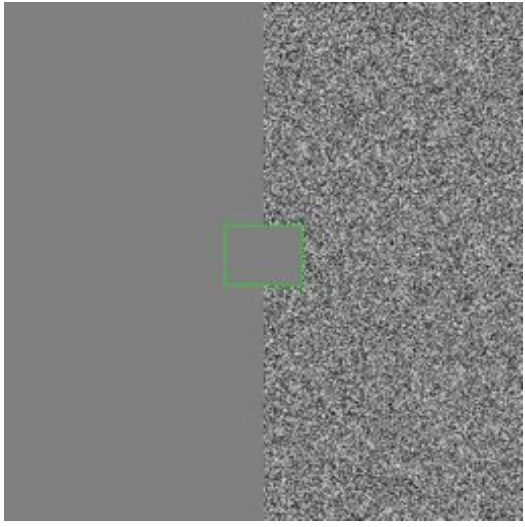
Image example

Inpainting with modified patch distance

Image example

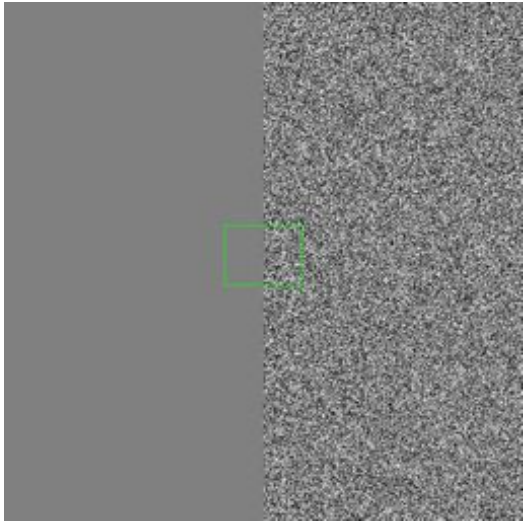
Original image

Noise example



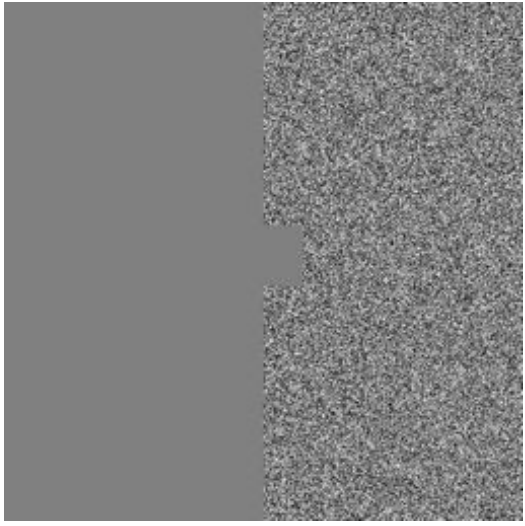
Inpainting with unmodified patch distance

Noise example



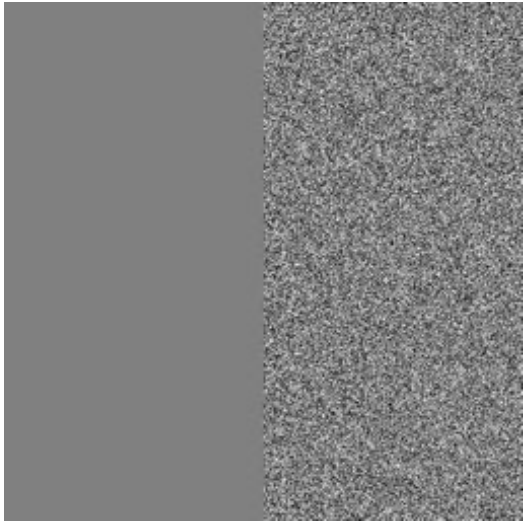
Inpainting with modified patch distance

Noise example



Inpainting with unmodified patch distance

Noise example



Inpainting with modified patch distance

Video Example – *original video*

Video Example - *unmodified patch distance*

Video Example - ***modified*** *patch distance*


Inpainted result – *spot the disocclusion!!*


Original video

More demos / paper / source code (to come) in

http://perso.enst.fr/~almansa/video_inpainting/