

Depth-based image completion for view synthesis

18th May 2011

Josselin Gautier, Olivier Le Meur,
Christine Guillemot

TEMICS team

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTEMES ALÉATOIRES



Outline

- Introduction and problematic
- Principle and approach
- Contributions
- Experiments
- Results
- Conclusion and perspectives

Introduction

- The virtual view synthesis is an hot topic for 3DTV applications:

- 2D-to-3D conversion

- Multi-view rendering for autostereoscopic displays

- Freeviewpoint TV



- Multi View + Depth (MVD) dataset : N texture + N depth maps



Problematic

- How to efficiently compress and render this huge dataset?
 - Take a subset of MVD views, e.g View 4, 5 & 6
 - Take a subset of views and interpolate the intermediate views:

View ₃



View ₅



View ₇



- Use only a **single** view and **extrapolate** the views → Our proposal with **inpainting** method

View _{5→4}



View ₅

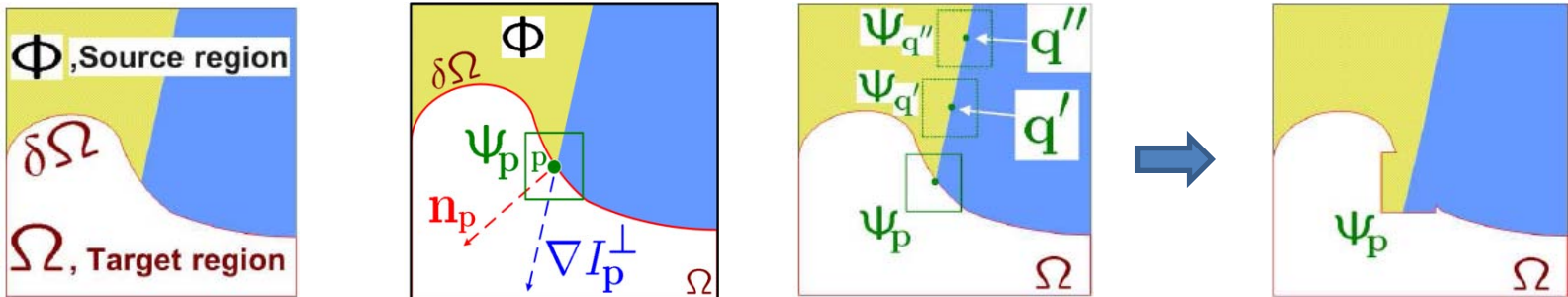


View _{5→6}



Inpainting principle

- Fill-in missing areas or artefacts on a 2D image (by diffusion or patch)
- Exemplar-based technique [Criminisi 04]:
 - Filling of disoccluded areas by best candidate patch sampling



◦ Two-pass process :

1. patch priority calculation $P(p) = C(p)D(p)$

$$C(p) = \frac{\sum_{q \in \Psi_p \cap (\mathcal{I} - \Omega)} C(q)}{|\Psi_p|}, \quad D(p) = \frac{|\nabla I_p^\perp \cdot n_p|}{\alpha}$$

2. Best candidate extraction and filling $\Psi_q = \arg \min_{\Psi_q \in \Phi} d(\Psi_p, \Psi_q)$

Approach

- Scene approach: different objects, a common background
- Depth can be known (or partially)



- Disocclusion holes appears at **specific** areas:

- at left or right of the foreground objects (previously occluding)
- belonging to the background plane
- ➔ proposal: fill-in the holes with **background texture only**

use the warping direction to ensure to begin with the background

Outline

- Problematic
- Inpainting principle
- **Contributions**
- Experiments
- Results
- Conclusion and Perspectives

Patch priority computation

Priority based on Criminisi algorithm:

$$P(p) = H [C(p)D(p)]$$

- $C(p)$ confidence term
- $D(p)$ tensor on RGBZ:

$$J = \sum_{l=R,G,B,Z} \nabla I_l \nabla I_l^T = \sum_{l=R,G,B,Z} \begin{pmatrix} \frac{\partial I_l^2}{\partial x} & \frac{\partial I_l}{\partial x} \frac{\partial I_l}{\partial y} \\ \frac{\partial I_l}{\partial x} \frac{\partial I_l}{\partial y} & \frac{\partial I_l^2}{\partial y} \end{pmatrix}$$

we use coherence norm [Weickert] based on eigenvalues

$$D(p) = \alpha + (1 - \alpha) \exp \left(\frac{-C}{(\lambda_1 - \lambda_2)^2} \right)$$

$$\lambda_1 \gg \lambda_2, D(p) \rightarrow 1 - \alpha$$

$$\lambda_1 \approx \lambda_2, D(p) \rightarrow \alpha$$

- $H [.]$ directional term depending on camera motion



Texture and structure propagation

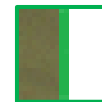
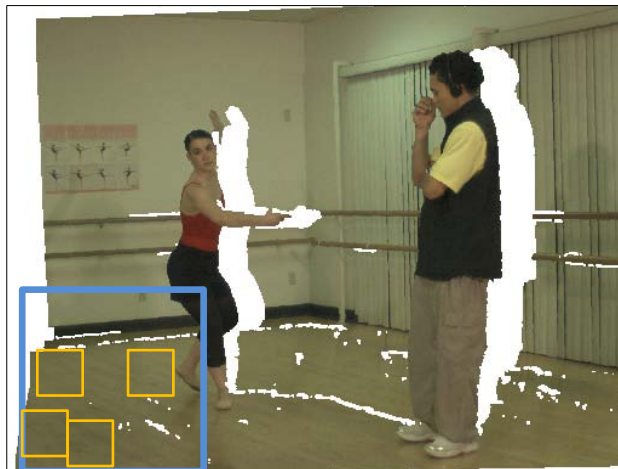
1. Search for K best candidates in the same color and depth range [Daribo 10]

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q)$$

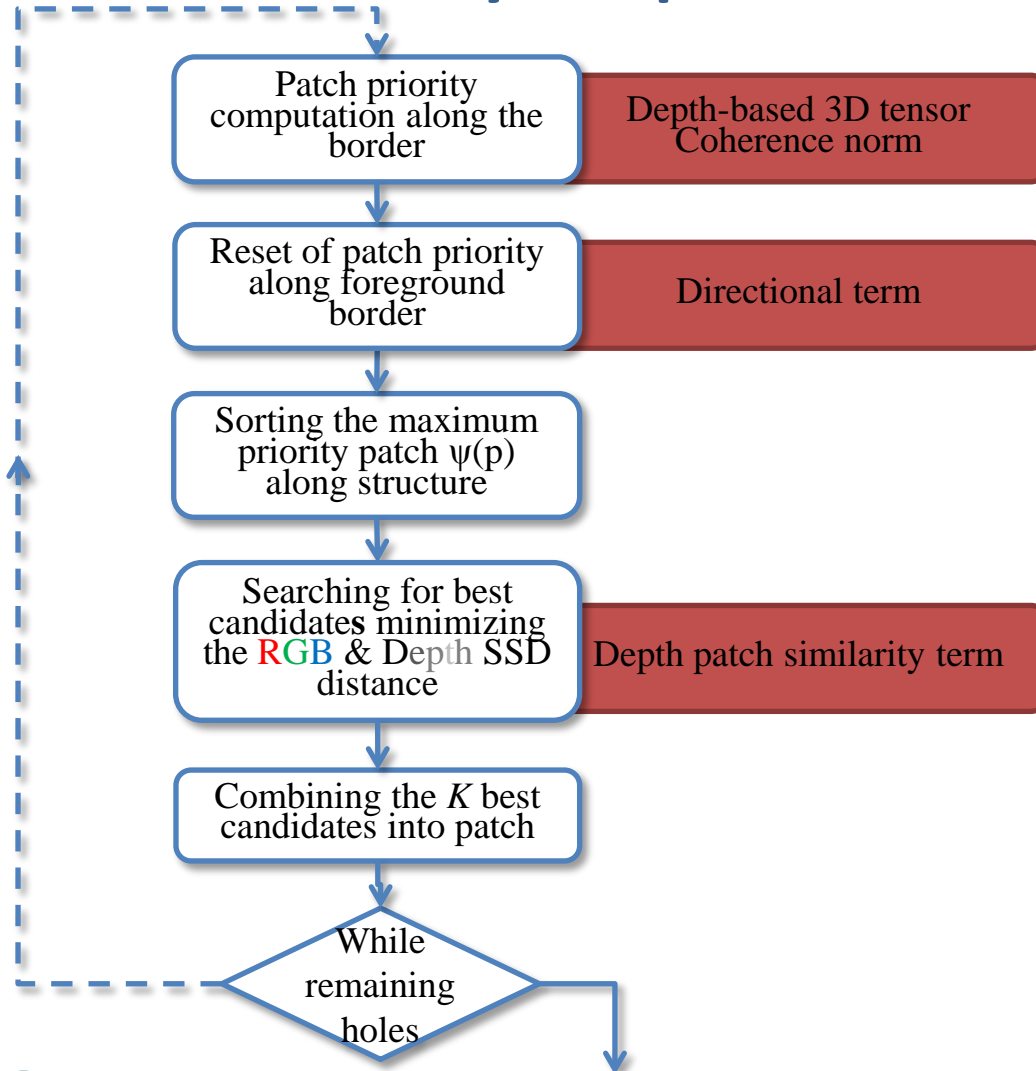
$$\text{with } d = \sum_{p,q \in \Psi_{p,q} \cap \Phi} \alpha_l \|\Psi_{\hat{p}} - \Psi_q\|^2$$

($l \in R, G, B, Z$ with $\alpha_{R,G,B} = 1$ and $\alpha_Z = 3$)

2. Exponential –based weighted combination of the K best candidates [Wexler 09]



Synoptic



Experiments

- Experiments on half resolution images (512x384) of Microsoft Research [1]
- Size of patch: 9x9, window: 49x49 pixels
- C Implementation of inpainting method (#5-10sec, fast but not optimized)
- Complexity varying with the image, patch and window size: $O(n^2)$
- Procedure:
 1. Projection (warping) of one original view to a virtual view
 2. Cracks filling and background anti-ghosting
 3. Inpainting method:
input: color image warped to new view +
ground truth depth map (in final view)
output: final inpainted view...



Results

- Ballet $V_5 \rightarrow 2$



Results for close view synthesis

Ballet $V_5 \rightarrow 4$



Criminisi
2D inpainting



Daribo
2D+Z inpainting



Our method
2D+Z inpainting

Results for distant view synthesis

Ballet $V_5 \rightarrow 2$



Criminisi
2D inpainting



Daribo
2D+Z inpainting



Our method
2D+Z inpainting

Results

Breakdancers $V_5 \rightarrow 2$



Criminisi
2D inpainting



Daribo
2D+Z inpainting



Our method
2D+Z inpainting

Application to freeviewpoint navigation

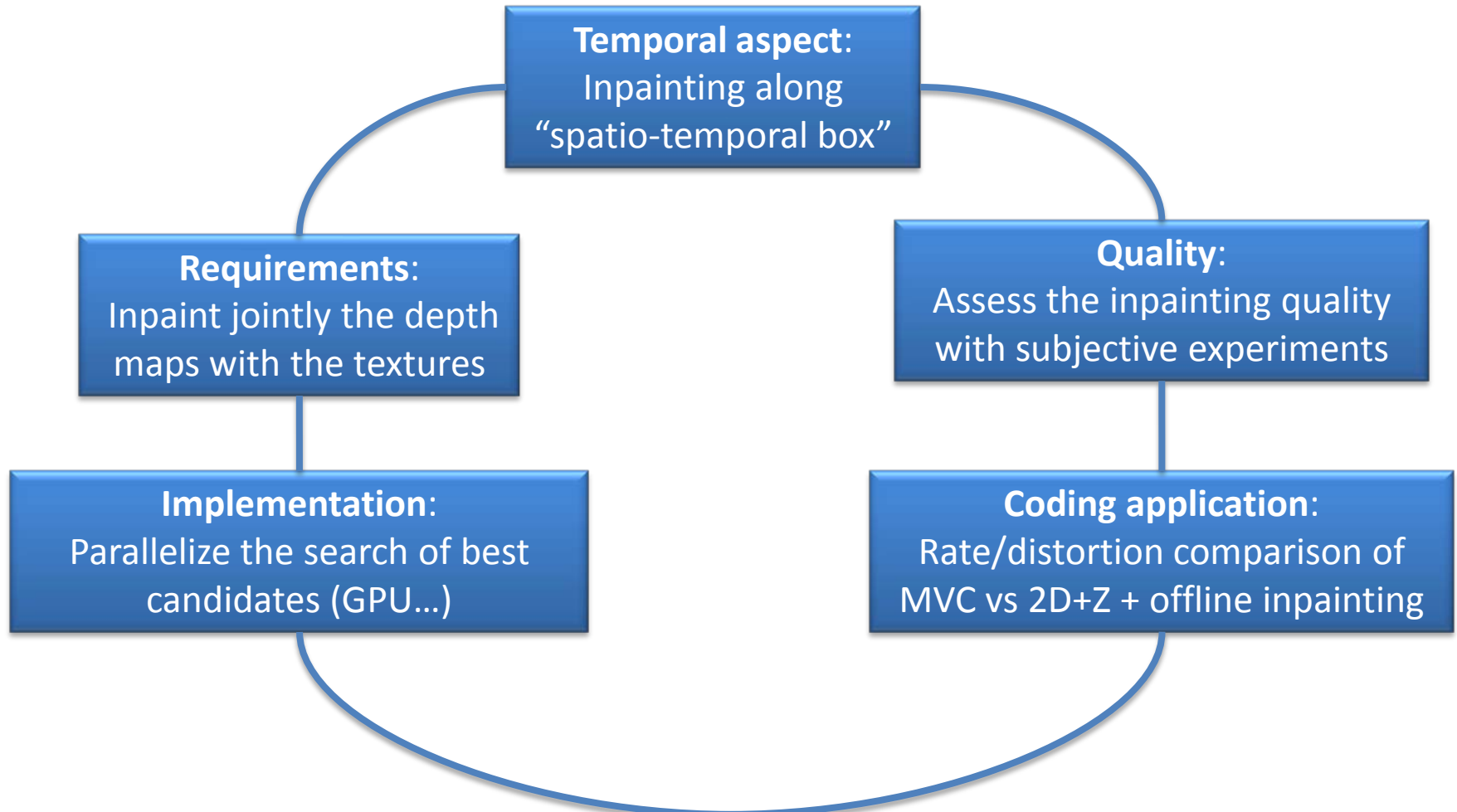
- Up to where can we navigate with depth-aided inpainting method?



Conclusion

- A new method of extrapolation of view based on structure and object in scene
 - Scene approach enable smart filling
 - Depth coherence terms reinforce the structure propagation
- Shows qualitative improvement on synthesized areas
- Suitable for close view extrapolation (on FTV dataset) to distant one (on 3DTV)
- Open issues:
 - Temporal coherence
 - Time i.e complexity growing with the square of window size and patch size
 - but search is highly parallelizable [Barnes 09]
 - No perceptual metric for relevant “best patches” selection
 - How to assess the quality of a inpainted view with its original one through an appropriate metric without including projection distortion? (registration?)

Perspectives



Thank you for your attention
Questions ?



References

- C. Barnes, E.Schechtman, A.finkelstein, “PatchMatch: a randomized correspondance algorithm for structural image editing”, ACM Transactions on Graphics (TOG), **2009**
- A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” IEEE Transactions on Image Processing, vol. 13, no. 9, pp. 1200–1212, **2004**.
- I. Daribo and B. Pesquet-Popescu, “Depth-aided image inpainting for novel view synthesis,” in IEEE International Workshop on Multimedia Signal Processing, **2010**.
- J. Weickert, “Coherence-enhancing diffusion filtering,” International Journal of Computer Vision, vol. 31, no. 2, pp. 111–127, **1999**.
- Y. Wexler, E. Shechtman, and M. Irani, “Space-time completion of video,” IEEE transactions on PAMI, pp. 463–476, **2007**.
- [1] C.L. Zitnick, S.B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, “High-quality video view interpolation using a layered representation,” in ACM SIGGRAPH. ACM New York, NY, USA, **2004**, pp. 600–608.