DEPTH GUIDED IMAGE COMPLETION FOR STRUCTURE AND TEXTURE SYNTHESIS

Michael Ciotta, Dimitrios Androutsos*

Department of Electrical & Computer Engineering
Ryerson University
350 Victoria Street
Toronto, Ontario,Canada.
{mciotta,dimitri}@ryerson.ca

ABSTRACT

This paper presents a method that separates the image completion process into structure and texture synthesis. A method is first introduced for completing the respective depth map through the use of a morphological diffusion-based operation, ensuring that structure is propagated smoothly across the unwanted region. These estimated depth values are then used to guide completion of the missing color image texture. A fast exemplar based PatchMatch approach is adopted and an extension of the coherence-based objective function introduced by Wexler et al. is used to complete the resulting texture in both the color and depth images. Finally, we provide experimental results to demonstrate the superiority of our method against others in a variety of different scenes.

Index Terms— Image inpainting, image completion, texture analysis, restoration, optimization

1. INTRODUCTION

Image completion, also commonly known as image inpainting, is one of the most elementary yet most challenging tasks in image processing and computer vision. It is the process of restoring missing or damaged areas in digital images with information from its surroundings in such a way that a visually plausible outcome is obtained. This task is used in many applications such as repairing damaged photographs, removal of unwanted objects/markings in photos, removing scratches and stains from deteriorated images and etc [1]. While much progress has been made in the past couple years, image completion still remains a challenging problem due to the higher level information needed from the surrounding scene to fill in the missing region. Inpainting algorithms can then be classified into two broad categories: diffusion-based inpainting and exemplar-based inpainting. The differences are in how information is chosen and propagated throughout the unknown region.

* Corresponding Author

Diffusion based inpainting techniques fill in missing regions through the use of a diffusion process i.e. by smoothly propagating colour information from the boundary in-towards the interior of the missing region. This diffusion process can be achieved by solving a high order, non-linear partial differential equation (PDE) and propagating colour information along isophotes [2, 3, 4]. Recently, Guo et al. [5] and Jawas et al.[6] avoid the computation of the high order PDE and perform this diffusion process through the use of a morphological erosion operation. In both cases, the filled-in colour varies smoothly to keep the continuity of distinct image areas and suffer in texture synthesis.

Exemplar based inpainting techniques fill in the missing regions by searching for an optimal exemplar pixel /patch located within the source regions, for each unknown pixel /patch in the target region. The problem can be further formulated as a discrete Markov Random field (MRF) optimization problem [7, 8, 9] or a patch searching global optimization problem [10, 11, 12, 13]. A combination of [12] and [13] is implemented as the content aware fill in Adobe Photoshop (CS5) [14] which is arguably considered the current state-of-the-art region filling algorithm in terms of visual quality and convergence speed.

Examplar-based inpainting techniques, on their own, perform well for texture synthesis but fail in recovering desired structure, since the optimal content can be found anywhere within the image. In attempting to constrain this search space using stereo depth maps, the recent work of He et al.[15] uses the PatchMatch algorithm with a constrained search space to only search within regions that are farther in depth than the object being removed, as this is most likely the background. This method performs well for simultaneous colour and depth texture synthesis but fails in the recovery of linking similar distinct areas across the hole.

To overcome this issue, our approach uses the RGBD matching approach from [15] but improves upon the work as follows: First, we use a morphological diffusion based approach similar to [6] to estimate the missing depth information in order to propagate distinct image areas across the

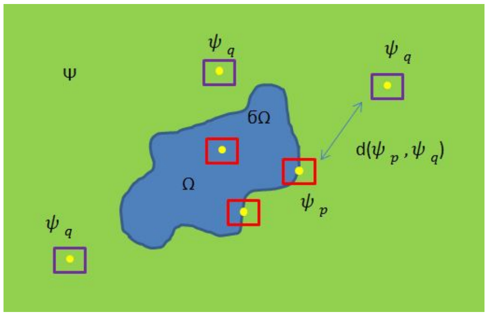


Fig. 1. Notation Diagram for exemplar-based inpainting

hole. Then, we iteratively optimize the synthesized texture using the diffused depth information as a constraint. Finally, we do not explicitly copy pixels from the known to unknown region during the colour/depth updating but adopt the weighting proposed in [12].

2. METHOD OVERVIEW

Given a colour image I and its corresponding depth map D, let Ω denote all missing pixels within I and D (target region) and $\Psi = \Omega^c$ be the known remaining pixels outside the missing region (source region) which are used to fill in the hole. Further let $\partial\Omega$ denote the boundary between the known and unknown regions. Using notation similar to [13] and [12], we seek to minimize the following measure of image coherence in both depth and colour images:

$$d_{total}(\Psi, \Omega) = \sum_{p \subset \Omega} \min_{q \subset \Psi} d(\psi_p, \psi_q), \tag{1}$$

where Ω is the target/missing region, Ψ is the source/known region, ψ_p is a square patch centered at $p \in \Omega$, ψ_q is a exemplar selection patch centered at $q \in \Psi$ and $d(\psi_p, \psi_q)$ is a measure of the Euclidean distance between patch ψ_p and ψ_q. By satisfying equation (1) we try to ensure that every patch located within the missing region (ψ_p) is filled with a known patch located within the known region (ψ_q), thus penalizing any unwanted artifacts and ensuring image coherence. The optimal exemplar we are searching for must satisfy

$$\psi_q = arg \min_{\psi_q \subset \Psi} d(\psi_p, \psi_q). \tag{2}$$

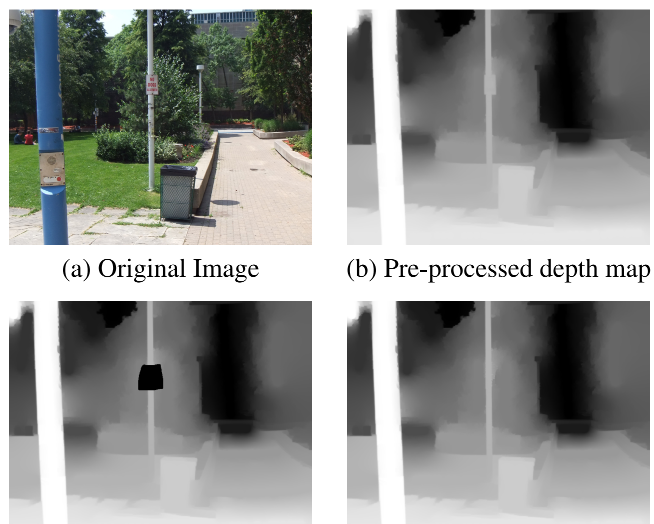
Figure 1 shows a graphical representation of the notation used for the filling algorithm.

2.1. Depth/Structure Completion

Unlike color images, which contain rich texture information, a depth map generally contains smooth regions with strong edges representing distinct object boundaries having defined depth relative to a desired viewpoint. Prior to use however, we first fill any occluded regions using a small hole diffusion based method propagating nearby depth values. After this pre-processing step, we use a method similar to [5] where, for each patch ψ_p centered on $p \in \Omega$, we iteratively erode and restore all the pixels on the current boundary of Ω through a structure/texture feature matching algorithm propagating known pixels into the unknown region. The algorithm proceeds as follows: First, within a search scope $\Psi_D = \Omega \oplus B$, where $\oplus$ denotes the morphological dilation operation, and B is a 3x3 structuring element, each known patch ψ_q with $q \in \Psi_D$ is compared to ψ_p using the distance measurement

$$d(\psi_p, \psi_q) = d_{depth}, \tag{3}$$

where d_{depth} is the Euclidean distance between ψ_p and ψ_q in depth. It is important to note that only valid pixels contained in Ψ are used in the computation of d_{depth} and we normalize d_{depth} by dividing by number of valid pixels. For each iteration, we repeat until each unknown pixel is flagged as known. Assuming distinct image areas are spatially continuous and are only separated by the hole, this algorithm provides diffusion from nearby pixels around the hole to be diffused into the hole propagating structure as shown in Figure 2 (d).



(a) Original Image (b) Pre-processed depth map

(c) Depth map with hole (d) Diffused Depth map

Fig. 2. Depth Inpainting Result

2.2. Texture and Colour Completion

We build directly on the PatchMatch-based approach of Barnes et al. [13] using an iterative update strategy, iteratively alternate between matching each patch ψ_p to its best match ψ_q for all $p \in \Omega$ and $q \in \Psi$ and then updating its colour. Our approach can be broken down into the following stages:

1. **Optimization** We define our objective function as:

$$d(\psi_p, \psi_q) = \alpha d_{colour} + \beta d_{depth}, \quad (4)$$

where d_{colour} and d_{depth} are the sum of absolute differences between ψ_p and ψ_q in colour and depth respectively, and α and β affect the contribution of d_{colour} and d_{depth} respectively. In our experiments we use α is 1.7 and β is 1.3 and we search for each ψ_q to satisfy (2).

2. **Initialization** Different from [13] and [15], we use the completed depth map to constrain our randomly selected exemplars. Let q_R be a randomly selected exemplar located within a search region around Ω. Furthermore, let d_p be the completed depth information corresponding to each $p \in \Omega$, then for each $p \in \Omega$ we select

$$p = \{q_R | q_R \in \Psi \cap q_R \in d_p\}. \quad (5)$$

Therefore, since our completed depth information contains completed structure information, we provide an educated guess corresponding to its desired texture by searching in a constrained area located around the missing region. This is in order to avoid searching in other objects that may have the same depth value.

3. **Propagation** For each overlapping patch $\Psi_p \in \Omega$ we compute the offset to the nearest neighbour patch $\Psi_q \in \Psi$. Our goal is to find the optimal offset map or approximate nearest neighbour field, i.e. the ANNF, for all $p \in \Omega$ ensuring (2) is satisfied. Let the coordinate of p be (x, y) and the offset between p and its exemplar q as $v(x, y)$. We attempt to reduce $D(v(x, y))$, where D is the distance function between two patches given in (4). On odd iterations, we attempt to improve $v(x, y)$ using the known offsets of $v(x-1, y)$ and $v(x, y-1)$ and taking the new value of $v(x, y)$ to be

$$\underset{(x,y)}{\operatorname{argmin}}(D(v(x, y)), D(v(x-1, y)), D(v(x, y-1))). \quad (6)$$

Alternatively, on even iteration we search in the alternate directions and attempt to improve $v(x, y)$ using the known offsets of $v(x+1, y)$ and $v(x, y+1)$ taking the new value of $v(x, y)$ to be

$$\underset{(x,y)}{\operatorname{argmin}}(D(v(x, y)), D(v(x+1, y)), D(v(x, y+1))). \quad (7)$$

Using the completed depth map as a constraint, we constrain the offsets search to only consider values containing the same completed depth information.

4. **Random Search** After each iteration, we attempt to improve $v(x, y)$ by testing a sequence of candidate offsets at an exponentially decreasing distance from $v(x, y)$. To avoid falling in suboptimal local minima, we only restrict the search range on the last iteration to ensure proper convergence.

2.2.1. Colour and Depth Updating

Let c be the colour and depth value of a target pixel $p \in \Omega$. After running several iterations of PatchMatch, we update the colour and depth value of pixel p using a weighted blending of values of the sources patches ψ_q matched to each target patch ψ_p that overlaps pixel p. The colour of pixel p is then given by

$$c = \frac{\sum w^i c^i}{\sum w^i}, \quad (8)$$

where c^i is the colour/depth value of pixel p given by a source patch ψ_q^i and w^i is a per-pixel weighting. We use the same distance transform weighting function as in [12], defined as

$$w^i = \gamma^{-dist(p,T)}, \quad (9)$$

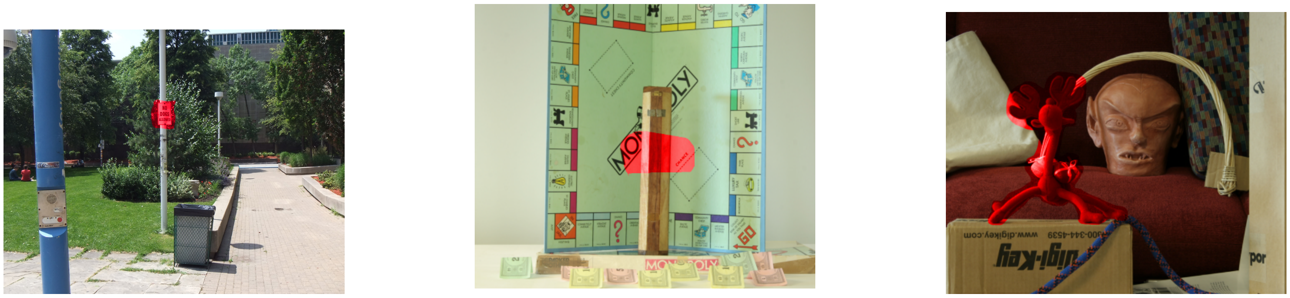
As proposed by Wexler et al. [12] we use γ = 1.3 and $dist(p, T)$ is the spatial distance from p to the boundary of the target region T.

3. RESULTS

Fig. 3 compares our method to the work published in [15] and [14]. As expected, the results produced from [15] and [14] draw textural content that is inappropriate for the desired structural completion (i.e. the wooden block in Fig. 3 middle and the pole in Fig. 3 left are not linked across the hole). Our results converge in 5 iterations for both colour and depth, do not suffer for these errors as the initially completed depth information is used to constrain the textural completion process, allowing similar regions to be linked across the hole with the desired texture. Additional Results can be found at `http://www.rnet.ryerson.ca/~mciotta/ICASSP2016/`

4. CONCLUSION

We have presented a method for image completion by diffusion based depth inpainting followed by exemplar-based texture synthesis. In contrast to other depth guided single image completion methods, our approach considers the structure propagation in allowing similar regions separated by the hole to be linked in the resulting depth map. Using the estimated completed depth, we then complete the colour image by constraining to the resulting depth values. For future work, we plan an implementing a self-adapting distance function given in (4) where constants α and β change according to the image.

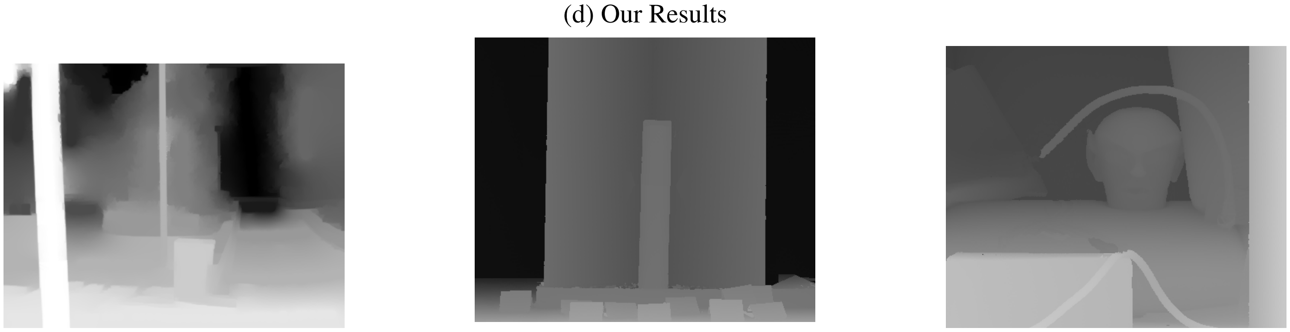


(a) Image with hole (red)

(b) He et al. [15] Results

(c) Content Aware Fill [14] Results

(d) Our Results



(e) Our Depth Results

Fig. 3. Colour and Depth Inpainting Result

5. REFERENCES

[1] C. Guillemot and O. Le Meur, "Image inpainting : Overview and recent advances," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 127–144, Jan 2014.

[2] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester, "Image inpainting," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 2000, SIGGRAPH '00, pp. 417–424, ACM Press/Addison-Wesley Publishing Co.

[3] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *Image Processing, IEEE Transactions on*, vol. 10, no. 8, pp. 1200–1211, Aug 2001.

[4] M. Bertalmio, A.L. Bertozzi, and G. Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2001, vol. 1, pp. I–355–I–362 vol.1.

[5] Hao Guo, N. Ono, and S. Sagayama, "A structure-synthesis image inpainting algorithm based on morphological erosion operation," in *Image and Signal Processing, 2008. CISP '08. Congress on*, May 2008, vol. 3, pp. 530–535.

[6] Naser Jawas and Nanik Suciati, "Image inpainting using erosion and dilation operation," in *International Journal of Advanced Science and Technology*, Feburary 2013, vol. 51.

[7] A.A. Efros and T.K. Leung, "Texture synthesis by non-parametric sampling," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1999, vol. 2, pp. 1033–1038 vol.2.

[8] Jian Sun, Lu Yuan, Jiaya Jia, and Heung-Yeung Shum, "Image completion with structure propagation," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 861–868, July 2005.

[9] N. Komodakis and G. Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *Image Processing, IEEE Transactions on*, vol. 16, no. 11, pp. 2649–2661, Nov 2007.

[10] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *Trans. Img. Proc.*, vol. 13, no. 9, pp. 1200–1212, Sept. 2004.

[11] N. Komodakis, "Image completion using global optimization," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, June 2006, vol. 1, pp. 442–452.

[12] Yonatan Wexler, Eli Shechtman, and Michal Irani, "Space-time completion of video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 463–476, Mar. 2007.

[13] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 24:1–24:11, July 2009.

[14] Connelly Barnes, Eli Shechtman, Ivan Belaunde, Dan B Goldman, and Jeff Chien, "Adobe content-aware fill," `http://www.adobe.com/technology/projects/content-aware-fill.html`, 2015, [Online; accessed 3-July-2015].

[15] Liu He, Michael Bleyer, and Margrit Gelautz, "Object removal by depth-guided inpainting," in *ÖAGM / AAPR Workshop 2011*, (2011), pp. 1–8.