Multiview Image Completion with Space Structure Propagation

Seung-Hwan Baek     Inchang Choi     Min H. Kim*
Korea Advanced Institute of Science and Technology (KAIST)
{shwbaek;inchangchoi;minhkim}@vclab.kaist.ac.kr

Abstract

We present a multiview image completion method that provides geometric consistency among different views by propagating space structures. Since a user specifies the region to be completed in one of multiview photographs casually taken in a scene, the proposed method enables us to complete the set of photographs with geometric consistency by creating or removing structures on the specified region. The proposed method incorporates photographs to estimate dense depth maps. We initially complete color as well as depth from a view, and then facilitate two stages of structure propagation and structure-guided completion. Structure propagation optimizes space topology in the scene across photographs, while structure-guide completion enhances, and completes local image structure of both depth and color in multiple photographs with structural coherence by searching nearest neighbor fields in relevant views. We demonstrate the effectiveness of the proposed method in completing multiview images.

1. Introduction

Digital photography has been stimulated from the perspective of collaboration, yielding casual multiview photographs of a scene. Even though multiview images are getting popular from the recent advances of digital imaging devices such as mobile cameras and light field cameras, consistent image editing of multiview photographs rarely has been discussed. For instance, most traditional image completion methods focus on an individual photograph [9, 31, 2]; hence, applying such single image completion methods is not able to retain the geometric consistency of space structures in images.

Traditional inpainting [17, 28] focuses on preserving coherency of structure in an image. In contrast, we propose a novel solution that allows for geometric coherency, so-called space structure, while completing multiview images. Multiview photographs allow us to estimate depth, utilized as geometric constraints. While completing multiview images, our method accounts for not only structure in an image, but also space structure across multiview images.

Our method first receives a target region from a user in one of the multiview photographs. The region is then completed by three steps: (1) the preprocessing step estimates dense depth maps and camera parameters using structure from motion, and completes the color and the depth of a reference view, (2) the structure propagation step conveys the space structure of the previous completions to the next image to be completed, and (3) the structure-guided completion enhances the quality of completion by incorporating the transferred space structure, searching for local similarity via the nearest neighbor field across multiview photographs. We perform the structure propagation and structure-guided completion steps for all images, resulting in multiple completed images for the first iteration. To enhance the image completion, our iterative approach repeats these two steps until we reach to the final completed images.

2. Related Work

Image completion synthesizes image structure to fill in larger missing areas using exemplar [9, 31, 2, 10, 17]. Our work is a branch of image completion, focusing on image completion of multiple photographs.

Single Image Completion Criminisi et al. [9] prioritized patches to be filled, and greedily propagated the patches from the known source region to the target region. Wexler et al. [31] solved the image completion problem with an expectation-maximization (EM) optimization using an image pyramid. Since the nearest neighbor field (NNF) search is the computational bottleneck of image completion, Barnes et al. accelerated the process by introducing an approximated search algorithm [2]. Darabi et al. [10] considered the gradients of images for patch-based synthesis and also introduced alpha-blending of multiple patches. Our method is based on the framework of Wexler et al. [31] with the main difference that our approach simultaneously completes the color and the depth of multiview photographs with structural coherence.
Stereo Image Completion Since the disparity of similar colors depends on depth, completing the color and the depth on a target region in stereo images is a chicken-and-egg problem. Wang et al. [30] jointly synthesized the color and the depth using a greedy search [9]. They executed the image completion independently on left and right images and then evaluated them to constrain the perspective consistency. However, the consistency of image completion is not always guaranteed, particularly when processing a large holes. Morse et al. [23] initially inpainted the depth via a simple interpolation with a PDE solver, and exploited the inpainted depth to complete the color. Since PDE-based inpainting of the depth is even more challenging than completing the color due to lack of clues, the inaccurate depth completion in this method tends to reduce the quality of image inpainting significantly. Similar to traditional image-based rendering methods [32], the goal of these stereo inpainting methods is merely to remove target foreground objects by filling in the occlusion region, which is caused by the objects, using patches from the background. Note that large-scale structures in image completion cannot be accounted for the previous stereo completion methods; e.g., the synthesized structure in one image could be outside of the target region on other images due to perspective projection. Alternatively, Luo et al. [22] asked users to manually complete the target depth map and then performed image completion on the color images only. While this method can complete the target region with foreground objects, users need to manually correct the target region of both views considering the visibility of completed depth.

Multiple Image Completion Hays and Efros [16] embedded scene semantics into target regions by using millions of prior Internet photographs. Darabi et al. [10] considered a mixture of two patches from different images in applying a random correspondence search method [2]. Barnes et al. [3] proposed a 3D data structure for efficient NNF search from multiple source images. Wexler et al. [31] introduced a video completion method that propagates image completion with a smoothness constraint of consecutive frames. Even though these methods accounted for multiple images or patches, they presumed that the camera view point does not change significantly. Multiple images are merely employed to improve the quality of image completion. They all disregard the three-dimensional space structure of image completion. To the best of our knowledge, our method is the first work that accounts for the geometric consistency of space structure in image completion of multiple photographs.

3. Multiple Image Completion

Our pipeline consists of three steps: preprocessing, structure propagation, and structure-guided completion. See Figure 1 for an overview.

Input We use multiview unstructured photographs $I_i$ containing a static scene from different views $v_i \in V$ as input, where $V$ is the set of every view. A user specifies an original image $I_o$, and give an original mask $M_o$ indicating the
completion region $\Omega_o$ to be completed on the original image $I_o$. Note that our method automatically completes the corresponding target region $\Omega_i$ on the $i$-th image, resulting in a consistent completed image $\hat{I}_i$.

### 3.1. Preprocessing

Since we target casually taken multiple photographs for image completion, preprocessing starts to estimate dense depth maps using structure from motion (SfM). See Figure 2 for example. We then conduct initial color and depth image completion on a reference view $v_o$. See Sections 3.3 and 4 for details on structure-guided image completion and SfM implementation.

### 3.2. Structure Propagation

Subsequent to estimating the camera parameters from SfM, we first select relevant views, where the initial completion is visible in multiple photographs. The completed color and depth in a target region of $v_o$ is propagated to guide the completions of the selected views to preserve geometric consistency. See Figure 3 for overview.

#### Inverse & Direct Projection

Let $\hat{V}$ be the set of completed multiviews, where $v_i \in \hat{V}$ is progressively completed from the original view $v_o \in \hat{V}$. In order to determine next view to be completed $v_j \notin \hat{V}$, we compare the lastly completed view $v_i$ and other remaining views $v_j \notin \hat{V}$ in terms of the Euclidean distance of the camera positions and the orientation difference.

Once the most similar view $v_j$ is selected, the completed color $\hat{I}_i$ and depth $\hat{d}_i$ of the homogeneous pixel coordinates $\hat{p}_i \in v_i$ are projected inversely to the homogenous global coordinates $p$ via perspective projections of each view $v_i \in \hat{V}$:

$$ p = \hat{D}_i(p_i)E_i^{-1}K_i^{-1}\hat{p}_i, \quad (1) $$

where $K$ and $E$ are the intrinsic and extrinsic parameters of each view $v_i$ in relation to $\hat{p}_i$ and $p$.

Now we can accumulate each completion $\hat{p}_i$ in the global coordinates of a 3D point $p$, thus project image completion of $\hat{p}_i$ to $\hat{p}_j$ in the next view $v_j$ directly:

$$ \hat{p}_j = K_jE_jp. \quad (2) $$

This mechanism of inverse and direct projection allows us to preserve geometric consistency in multiple photographs although this results in cracks due to perspective projection and occlusion. These cracks are handled in the stage of structure-guided completion (see Section 3.3).

#### Identifying Completion Regions

Traditional stereo completion methods [30, 23] complete target region using background only; therefore, no additive structures are allowed in completion. In contrast, our structure propagation allows for introduction of additive structures in completion. To do so, we separate the target region $\Omega_j$ into two associated regions of addition and subtraction, $\Omega_j^+$ and $\Omega_j^-$ in the new view $v_j$. We need two steps to identify the complete regions in the new view. First, we should identify the subtractive regions $\Omega_j^-$ to be removed in $v_j$. Inversely project each pixel $\hat{p}_j \in v_j$ to corresponding $\hat{p}_o \in v_o$ using Equations (1) and (2). We then determine if the corresponding pixel $\hat{p}_o$ belongs to the original region $\Omega_o$. See Figures 4(a) and (b).

Second, we identify the additive regions $\Omega_j^+$ where some additive structure to be introduced in completion. Note that this $\Omega_j^+$ could be different from $\Omega_j^-$ depending on which structure is completed. Since the space structure exists only in the original view $v_o$, we project the completed pixel region $\Omega_o$ to $v_j$ via inverse and direct projection. This $\Omega_j^+$ therefore results in cracks and discontinuity. We expand the region to a convex hull that surrounds $\Omega_j^+$ to solve these artifacts using patch-based synthesis (see Section 3.3). Finally, we define a union region of these two regions as follows: $\Omega_j = \text{Conv}(\Omega_j^+ \cup \Omega_j^-)$.

Note that this process does not account for a new structure to be completed in $\Omega_j^-$ in the next step. We solve this problem by iterating this completion workflow rounding the series of photographs.

Compared to DIBR methods [24, 11, 14, 7], we synthesize missing regions as well as structure-guide regions in or-
3.3. Structure-Guided Multiple Completion

3.3.1 Space Structures

Contrary to state-of-the-art methods [31, 2, 10], we have additive completion regions \( \Omega_j^+ \) in addition to subtractive completion regions \( \Omega_j^- \). Prior to synthesizing the completion regions \( \Omega_j \), we first transfer the pre-completed color \( I_j \) and depth \( D_j \) on \( \Omega_i \) to \( \Omega_j^+ \) using Equations (1) and (2) as structure guide. We utilize this transferred information while synthesizing \( \Omega_j \) as guide. We denote the structure guides of color and depth as \( I_j^+ \) and \( D_j^+ \), respectively.

The traditional patch-based completion frameworks [31, 2] build an image pyramid that includes the target regions from coarse to fine levels. These methods search for a similar patch in the source region for the target patch centered at target pixel \( p_j \), in order to substitute the color of \( p_j \) with that of the similar source pixel, so-called search. Now we have a source patch matched with target pixel \( p_j \), as source patches corresponding to the neighbor pixels of \( p_j \) could have pixels that lie on the pixel position \( p_j \), they interpolate the colors of those pixels to determine the final color of \( p_j \), so-called vote. For each level, search and vote are performed iteratively, and this approach results in a completed image at the finest level.

Our proposed approach imposes the structure guides through this optimization problem. We synthesize the target region \( \Omega_j \) while maintaining the projected structure guides of \( I_j^+ \) and \( D_j^+ \) on \( \Omega_j^+ \). Our intuition is that the structure guide is roughly correct while the details are crisp (see Figure 5). Our approach harmonizes patch-based synthesis with space-structure guide in order to achieve geometric consistency in multi-image completion.

3.3.2 Coherence vs. Structure Guide

Our goal is to synthesize the target region \( \Omega_j \) using the source regions of every view \( \Psi \) \( = \{ \Psi_k = A_k \setminus \Omega_k^{-1} | k \in V \} \), where \( A_k \) is the set of every pixel in view \( k \). Our energy function consists of two terms, a coherence and a guidance term:

\[
E = \sum_{p_j \in \Omega_j} \min_{q_k \in \Psi_k} \{ E_{coherence}(p_j, q_k) + E_{guide}(p_j) \},
\]

where \( p_j \) is a target pixel on the target region \( \Omega_j \), and \( q_k \) is a source pixel on the source region \( \Psi_k \).

The coherence term \( E_{coherence} \) describes the similarity of the target and source patch in terms of color in CIELAB [8] and depth gradients, defined by:

\[
E_{coherence}(p_j, q_k) = C \left( \bar{I}_j(p_j), \bar{I}_k(q_k) \right) + \alpha C \left( \nabla D_j(p_j), \nabla D_k(q_k) \right),
\]

where \( C \) is the L2 norm of the difference of two values, and \( \alpha \) is the weighting constant. \( \bar{I}_j(p_j), \nabla D_j(p_j) \) are the color patches and the depth gradient patches centered at pixel \( p_j \).

Note that this function resembles that one from Darabi et al. [10] with the following difference. While Darabi et al. took both color and color gradients in their coherence term, we take both color and depth gradients for evaluating space-structure coherence. Our principal intuition for examining depth gradients is that, if we evaluate the difference of the depths directly, we cannot utilize many patches of the same shape but in a different distance to complete a depth map. Our light-weight shape representation of depth gradients allows us to evaluate the shape similarity of depth patches to synthesize depth information in missing regions, whereas traditional stereo completion methods [30] examined direct depth information. \( \alpha \) varies depending on the level of quality of the reconstructed depth maps.
The guidance term $E_{\text{guide}}$ accounts for the similarity of the target patch and the structure guide for pixel $p_j \in \Omega_j^+$:

$$E_{\text{guide}}(p_j) = \beta C(I_j(p_j), I_j^+(p_j)) + \gamma C(D_j(p_j), D_j^+(p_j)), \quad (4)$$

where $\beta$ and $\gamma$ are the weighting factors for color and depth differences, respectively. The overall scales of $\beta$ and $\gamma$ balance the local image structure and the space structure guide. In order to solve this objective function, we complete the color image first by our patch-based completion considering structure guide $I_j^+$. We then complete the depth using the completed color as well as the structure guide $D_j^+$. This guide energy term is mainly handled by the guided voting stage. See Section 3.3.3 for detail on optimization.

### 3.3.3 Color and Depth Optimization

Our optimization is originated from Darabi et al. [31] with the main difference for structure-guided image completion. Our optimization consists of two steps: search and guided vote. Here we mainly describe the differences from the base algorithm.

**Guide Initialization** We first construct color and depth image pyramids of the original data and the structure-guide, $I_{j,i}, D_{j,i}$ and $I_{j,i}^+, D_{j,i}^+$, respectively, using a Lanczos kernel following Darabi et al. [10]. Here let $l \in [1, L]$ be the level of the pyramid and $L$ be the number of levels ($I_{j,L} = I_j$). Accordingly, target completion region $\Omega_j$ and sub-region with structure guide $\Omega^+_j$ are transformed into each level $l$, resulting in $\Omega_{j,i}^l$ and $\Omega^+_j$. Then, the structure guide $I_{j,i}^+$ at the coarsest level is initially copied to the additive completion region $\Omega^+_j \in I_{j,i}$ to impose the coarse image completion resemble the structure guide. The rest of completion region $\Omega_j \setminus \Omega^+_j$ is filled in by interpolating the gradients of the boundary pixels with a PDE solver. See Algorithm 1 for an overview.

**Search** For this search stage, we mainly use Equation (3) to find the NNF of the closest patch at pixel $p_j \in \Psi_p$ to match patch at pixel $p_j \in \Omega_j^+$. Here we are motivated to adopt a popular search method, proposed by Barnes et al. [2], which is based on random search and propagation. Since our method takes multiple images as input, we extend the search range from a single to multiple photographs, by concatenating them as one, similar to Barnes et al. [3].

**Guided Vote** Our guided vote is the main difference from the state-of-the-art methods. From the search process, we collect colors and depth gradients of the patches from the NNF of neighbor pixels. We estimate initial color and depth gradients from the NNF, $I_{j,i}^+(p_j)$ and $D_{j,i}^+(p_j)$, centered at target pixel $p_j$ using the weighted sum approach, proposed by Wexler et al. [31]. These coherent color and depth values are updated with a guidance of space structure, as described in Equations (3) and (4), as follows.

Since we obtain the initial colors $I_{j,i}^+$ from coherence search, to account for space structure, we update color at pixel $p_j$ by linearly interpolating the coherent color $I_{j,i}^+(p_j)$ and structure guide $I_{j,i}^+(p_j)$, where $p_j$ is on the additive completion region $\Omega_j^+$:

$$\tilde{I}_{j,i}(p_j) \leftarrow (1 - \beta(l)) \times I_{j,i}^+(p_j) + \beta(l) \times I_{j,i}^+(p_j), \quad (5)$$

where $\beta(l)$ is the frequency-dependent weight. We define $\beta(l)$ as a truncated function that interpolates the level of weighting linearly:

$$\beta(l) = \begin{cases} 
1, & l \leq l_s \\
(l_e - l)/(l_e - l_s), & l_s < l \leq l_e \\
0, & l_e < l 
\end{cases}, \quad (6)$$

![Figure 6](image_url)

Figure 6. (a) shows an initial completion of the original view using our RGB-D completion method. (b) presents our intermediate projection of the completed structure to a new view. Note that the distorted space structure is caused by inaccurate depth. (c) shows the result of our completion. The distorted color and depth are updated iteratively through patch-based synthesis with structure guide. (d) compares our result with a traditional DIBR method, which complete the missing region with background and is incapable of handling depth inaccuracy.

---

**Algorithm 1 COMPLETION()**

**Input:** $I_{j,i}, I_{j,i}^+, D_{j,i}, D_{j,i}^+$

**Output:** $\tilde{I}_{j,i}, \tilde{D}_{j,i}$

1: Create image pyramids $I_{j,i}, I_{j,i}^+, D_{j,i}, D_{j,i}^+$
2: Initialize $\tilde{I}_{j,1}, \tilde{D}_{j,1}$
3: for $l = [1, ..., L]$ do
4:  $\text{NNF}_{j,i} \leftarrow \text{SEARCH}(\tilde{I}_{j,i}, I_{j,i}^+, \nabla \tilde{D}_{j,i}, \nabla D_{j,i}^+)$
5:  $\tilde{I}_{j,i}, \tilde{D}_{j,i} \leftarrow \text{VOTE}(\text{NNF}_{j,i}, I_{j,i}^+, I_{j,i}^+, \nabla D_{j,i}^+, D_{j,i}^+)$
6: end for
where \( l_s \) and \( l_c \) are the user parameters that control the influence of structure guide \( I_i' \).

Although we compare depth gradients in NNF search, our goal is to complete depth \( \hat{D}_{j,l} \) rather than \( \nabla D_{j,l}^* \) on the target region \( \Omega_{j,l} \) and to achieve geometric consistency with structure guide \( D_{j,l}^* \). For the region \( \Omega_{j,l}^c = \Omega_{j,l} \setminus \Omega_{j,l}^+ \), we want to complete depth \( \hat{D}_{j,l} \) from the voted depth gradients \( \nabla D_{j,l}^* \). Also, we should preserve space structure \( D_{j,l}' \) projected from the previous view, for the other region. Finally, the optimal \( \hat{D}_{j,l} \) can be estimated by minimizing the following objective function:

\[
\min_{\hat{D}_{j,l}} \sum_{\Omega_{j,l}^c} \| \nabla \hat{D}_{j,l} - \nabla D_{j,l}^* \|^2, \tag{7}
\]

with boundary conditions \( \hat{D}_{j,l}|_{\partial \Omega_{j,l}^c} = D_{j,l}'|_{\partial \Omega_{j,l}^c} \). We optimize this objective equation with a constraint term of the boundary conditions: \( \gamma \left( \hat{D}_{j,l}|_{\partial \Omega_{j,l}^c} - D_{j,l}'|_{\partial \Omega_{j,l}^c} \right)^2 \), where a constant \( \gamma \) is a weighting constant that appears in Equation (4). By applying the Euler-Lagrange equation, the optimal \( \hat{D}_{j,l} \) satisfies the Poisson equation: \( \nabla^2 \hat{D}_{j,l} = \nabla \cdot D_{j,l}^* \).

Algorithm 2 summarizes our completion method.

### 4. Implementation Details

**Preprocessing** Point clouds and camera parameters are estimated by using SfM methods [27, 13] from multiple photographs. We then project the point clouds into each view, resulting in a sparse depth map per a view using Equation (1) (see Figure 2b). We propagate sparse depth estimates to every pixel, yielding a dense depth map \( D_i \) per view (Figure 2c). Once we build depth information in photographs, for the target region \( \Omega_o \) of input view \( v_o \), we perform our completion method (Section 3.3) without any structure guide, resulting in a completed color image \( \hat{I}_o \) and the completed depth map \( \hat{D}_o \).

**Multiple Source Images** We complete each view of multiple photographs with consideration of space structure, as described in Section 3. Different from traditional completion methods, we use multiple images as input, and therefore we extend the patch search range from a single to multiple photographs, by concatenating them as one following Barnes et al. [3]. When we complete view \( v_j \), we select \( M \) source views uniformly on the sorted views w.r.t. the pose difference with view \( v_j \), for computational efficiency. The concatenated image is used for our structure-guide completion (Section 3.3), whereas previous stereo completion methods [30, 22, 23] utilize one of the left and right images to complete the selected image.

**Depth Denoising** In general, a depth map reconstructed from the Poisson equation suffers from severe noise. We therefore conduct an additional process of depth denoising that smooths out the estimated depth in the target region. We first segment the initially-estimated depth map into superpixels [1] by using the color information in the completed color image. We then perform median filtering per each depth superpixel to clean noise. Matting Laplacian [20] is applied to remove quantization error over superpixel edges in the depth map.

**Parameters** We use a \( 7 \times 7 \) patch for color and depth. Parameter \( \alpha \) in Equation (3) controls the balance between color and depth gradients in calculating individual coherence within each image. \( \alpha \) is set to a lower value to weight the color similarity for local coherence of each image. Parameter \( \beta \) in Equation (4) is determined by \( l_s \) and \( l_c \). In Equation (6), \( l_s \) and \( l_c \) are the weights to enhance the depth errors more, which is originated from Poisson reconstruction. In this experiment, we set the values of \( \alpha \), \( l_s \), \( l_c \), and \( \beta \) as 0.01, 0.3, 0.6, and 100. See Figure 9. In addition, we use ten levels of the image pyramid and the number of source views is set to three. We found that two iterations are enough for satisfying geometric consistency across multiple views. We complete three to six images selected from three to twenty images used for SfM, considering computational feasibility.

### 5. Results

Our Matlab and C++ implementation took approx. 300 seconds to complete color and depth of a \( 680 \times 900 \) image with \( 2.5 \times 10^4 \) target pixels. Computing time increases linearly proportional to the number of input images in general, taking 25 minutes for the multiple images with five photographs. To validate the performance of our method, we applied our method to three types of input data: stereoscopic images, a single RGB-D image and multiple images.
Figure 7. Stereo image completion. Our method completes the stereo image (a) with complex structures, resulting in the completed images (b) and (c) with geometric consistency by creating or removing objects. Wang et al. \[30\] utilize background patches only in turn. They cannot preserve the structure of the scene, such as the pipe in the left-side of the scene (d) and (e). Also note that there are some notable artifacts in (d) and (e), which are originated from the individual completion of each view. Image courtesy of Scharstein et al. [26].

Figure 8. For a single RGB-D image data (a), our method completes the target region (red area), resulting in a consistent RGB-D image (b). Image courtesy of Kim et al. [18].

Figure 9. (a) We investigate the effects of parameters: (b) $\alpha$ (orange box), (c) $\beta(l_s, l_e)$ (purple) and (d) $\gamma$ (yellow). We fix other parameters while varying only one of them, with the following values: $\alpha = 0.01, l_s = 0.3, l_e = 0.6$ and $\gamma = 100$.

Figure 8 compares our method with Wang et al. \[30\] for a stereo image. Our method reconstructs the left and the right image with geometric consistency while completing the connection of left-most pipe naturally, while Wang et al. \[30\] fills the target region with background objects, inhibiting natural completion of the complex structure.

Figure 8 shows the result of completing color and depth in a single RGB-D image. Our method can achieve geometric consistency of not only color but also depth. Note that depth patches cannot be compared directly for shape similarity due to depth value differences. Our gradient-based depth synthesis allows us to compare shape similarity in terms of gradients through the NNF search.

To the best of our knowledge, multiple image completion has not been studied before. For evaluation, we therefore applied a stereo completion algorithm of Wang et al. \[30\] on a pair of stereo images consequently to process multiple images. Wang et al. \[30\] impose geometric consistency by comparing the colors of corresponding pixels in completed left and right images. Figure 10 demonstrates that our method can synthesize new structures consistently across multiple photographs, while maintaining coherency of each image. In contrast, Wang et al. \[30\] cannot account for large structures and fail to preserve both geometric consistency and coherence inside each image.

6. Discussion and Conclusions

Propagated depth maps could be unstable due to the errors originated from SfM (see Figure 6b). To solve this problem, we account for color and depth coherency in an image. The propagated space structure is simultaneously updated through iteration with consideration of spatial frequency. Nevertheless, the quality of depth maps could be degraded in case of dynamic or feature-less scenes, which potentially cause the failure of our method. Note that initial completion of the target image can be conducted by any other state-of-the-art single image completion methods \[17, 10\].

We have presented a novel multiple image completion method that preserves the geometric consistency among different views. The proposed method consists of structure propagation and structure-guided completion. Our structure-guided completion, which is designed as a single optimization framework, exhibits superior results in terms of coherency and consistency. Our versatile algorithm enables to complete not only multiple images, but also stereoscopic images. In addition it allows to fill the empty region with foreground as well as background objects, which has been challenging so far in the previous stereo completion methods \[30, 23\].

Acknowledgements

Min H. Kim gratefully acknowledges Korea NRF grants (2013R1A1A1010165 and 2013M3A6A6073718) and additional support by Samsung Electronics (G01140381) and an ICT R&D program of MSIP/IITP (10041313) in addition to Sung-Hyun Choi (Samsung) for helpful comments.
Figure 10. Multi-image completion. The first column shows input color and depth. Red color indicates completion regions. The second and the third column shows completed depth and color, respectively, compared with Wang et al. [30]. The fourth and the fifth column presents completions from different views. Refer to the supplemental materials for more results. Image courtesy of Olsson and Enqvist [25].
References


