Laplacian Patch-Based Image Synthesis

Joo Ho Lee Inchang Choi Min H. Kim*
Korea Advanced Institute of Science and Technology (KAIST)
{jhlee;inchangchoi;minhkim}@vclab.kaist.ac.kr

Abstract

Patch-based image synthesis has been enriched with global optimization on the image pyramid. Successively, the gradient-based synthesis has improved structural coherence and details. However, the gradient operator is directional and inconsistent and requires computing multiple operators. It also introduces a significantly heavy computational burden to solve the Poisson equation that often accompanies artifacts in non-integrable gradient fields. In this paper, we propose a patch-based synthesis using a Laplacian pyramid to improve searching correspondence with enhanced awareness of edge structures. Contrary to the gradient operators, the Laplacian pyramid has the advantage of being isotropic in detecting changes to provide more consistent performance in decomposing the base structure and the detailed localization. Furthermore, it does not require heavy computation as it employs approximation by the differences of Gaussians. We examine the potentials of the Laplacian pyramid for enhanced edge-aware correspondence search. We demonstrate the effectiveness of the Laplacian-based approach over the state-of-the-art patch-based image synthesis methods.

1. Introduction

In digital photography, we often confront a situation where certain causes, such as blocks by uninvited objects, occlusions, failures in transmission, and holes produced by different perspectives in binocular stereo, corrupt a portion of images. Accordingly we may wish to fix these corruptions with plausible contents. To address these situations, *image inpainting* [6], which refers to filling in a corrupted area, has received attention during the past decade. Although image inpainting has been commonly used even with novice users, there is no completely versatile inpainting algorithm that gives robust results under any conditions.

The image gradients are employed in classical inpainting methods [10] for detecting and copying image structure from the boundary to the interior such that the image gradients, the first derivatives of image intensity, measure directional changes of intensity around edges. The *image pyramid*, which denotes an image representation based on multiscale signals, has been widely used as a typical practice for enhancing *structural coherence* when completing missing regions in patch-based synthesis [29]. Recently, the *image gradients* in each level of the *image pyramid* are used for enhancing edge structure in addition to coherent patch-based synthesis [11]. Even though the latest approach of combining gradients and the image pyramid has improved the structural coherence and details in inpainting, the gradient operator is directional and thus requires twofold greater computation of multiple operators, and it introduces a significantly heavy computational burden to solve the Poisson equation. Furthermore, it presents inconsistency often with artifacts in non-integrable gradient fields.

The *Laplacian* operator, which is the divergence of gradients of image intensity, takes advantage of being isotropic and invariant to rotation (Figure 1c). In addition, coordinates of the Laplacian correspond to those of the edges, being well aligned to represent the image structure over edges. The *Laplacian pyramid* allows us to decompose the *base* and *detail structure* of an image into different spatial frequency components that can preserve structure upon decomposition. This representation has been used in many applications such as image blend/fusion, enhancement, and denoising. However, to the best of our knowledge, the potentials of the Laplacian pyramid has not been intensively exploited in previous coherent patch-based image synthesis. The proposed method is the first work that combines the *Laplacian* with patch-based synthesis of *global coherence*. In this paper, we examine the properties of the Laplacian pyramid for image completion and describe our edge-aware patch-based synthesis using a Laplacian pyramid.

2. Related Work

Proposed first by Bertalmio et al. [6], inpainting refers to the task of filling in or completing holes, or missing or corrupted regions in images. Inpainting is classified into two categories: diffusion-based and exemplar-based methods.

*Corresponding author

Diffusion-Based. Methods that belong to this category complete missing areas by propagating the geometric structures of the neighboring areas. This is accompanied with the smoothness constraints that enforce the connectivity of local structures. Diffusion can be performed locally by solving a partial different equation (PDE) [22, 26, 27] and there is also a global optimization approach that minimizes the total variation [23] of the inpainted area [2, 24, 17]. Diffusion-based inpainting is shown to be very effective in reconstructing line, curves, and small holes, but it suffers from blurring artifacts when completing large holes.

Exemplar-Based. To extend the coverage of diffusion-based inpainting, exemplar-based approaches have been proposed. Within the unknown hole area this approach first determines the order, in which target patches are filled in, and then it searches similar candidate patches from known areas. Finally it composites the candidate patches on the location of the target. Subsequent to the approaches proposed by Criminisi et al. [10] and Drori et al. [12], and then succeeded by the global optimization of Wexler et al. [29], variants of these works have been resorted. Sun et al. [25] propagated the structures along with user-provided-guidance. Buyssens et al. [8] addressed the filling order using a tensor-based data term. The candidate patches, being represented by the nearest neighbor field (NNF), were searched following the nonlocal denoising algorithm [7]. This part was accelerated by introducing versatile data structures such as kd-trees [5] and vp-trees [31]. Barnes et el. [4] proposed a randomized search method that made computation much more tractable, thus it has been broadly adopted to search correspondence. In addition, Komodakis and Tziritas [16] proposed priority belief propagation to address patch composition via discrete global optimization. He and Sun [13] showed that exploiting the statistics of patch offsets is effective when searching the candidate patches, and Huang et al. [14] attempted to use planar structure guidance in order to take the perspective projection into account. Additionally, there have been efforts to adopt the patch sparsity [30], or to take the image super-resolution algorithm [19].

Further Coherence. Traditional patch-based inpainting algorithms, originally proposed by Wexler et al. [29, 3, 4], which compute similarity according to colors on the image pyramid, generated visually plausible results. However, some results have an inconsistent structure due to a lack of spatial coherence. Recently, the image gradients in each level of the image pyramid have been used for enhancing structural coherence [11]. However, Darabi et al. inherited the natural limitation of the image gradient operators. Since the gradient operator is directional, they compute horizontal and vertical changes of intensity, and the gradients must be solved by the Poisson equation for integration. Furthermore, this PDE-based solution often introduces artifacts in non-integrable gradient fields. In addition, Kalantari et al. [15] enhanced the patch-based searching algorithm with additional masks that account for foreground and background, resolving typical artifacts that occurs by the traditional patch-based synthesis.

Even though both *patch-based synthesis* and *the Laplacian pyramid* have been practiced extensively for recent decades, only a few works have examined the combined application of these two approaches. To the best of our knowledge, only Drori et al. [12] and Padmavathi and Soman [20] utilized the Laplacian approach to solve the inpainting problem. However, the both methods take the classic heuristic inpainting approach, originally proposed by Criminisi et al. [10]. Different from Criminisi, Drori et al. [12] use the Gaussian pyramid for coarse-to-fine refinement and adaptive patch size depending on texture complexity in addition to the usage of the Laplacian pyramid to naturally blend the target and source patch. Padmavathi and Soman [20] fundamentally follow Criminisi's heuristic approach. Even though they utilize the Gaussian and the Laplacian pyramid to separate the texture and the structure, they merely transfer the upsampled detail of low frequency to the high frequency layer without reconstructing details in the Laplacian domain. These two methods do not fully exploit the advantages of the Laplacian pyramid with the benefit of patch-based global optimization [29] and also inherit the limitations of the heuristic approach [10]. Furthermore, they suffer from global inconsistency, resulting in a coherent local decision around image structures. In contrast, our Laplacian approach is built on the state-of-the-art work of global energy optimization proposed by Wexler et al. [29] to overcome the incoherence problem and it enhances the structural coherence by taking advantages of the Laplacian representation.

3. Laplacian Patch-Based Image Synthesis

Our inpainting leverages a Laplacian pyramid to improve *structural coherence* in image synthesis. Our nearest neighbor search on a Laplacian pyramid is more invariant to rotation with propagated structural information and therefore we can improve the accuracy of the correspondence search compared with the state-of-the-art methods [29, 3, 11] do. Also, by making use of the upsampled Gaussian and Laplacian images, we obtain more robust performance against the noise and the change of parameters. This section provides the details of proposed edge-aware image synthesis strategy.

3.1. Laplacian Coherent Spaces

Gaussian, Gradient and Laplacian. The image pyramid offers a multi-resolution representation of an image [1] and has been practiced in many applications. Creating a pyramid consists of two steps: *filtering* and *sampling*. As

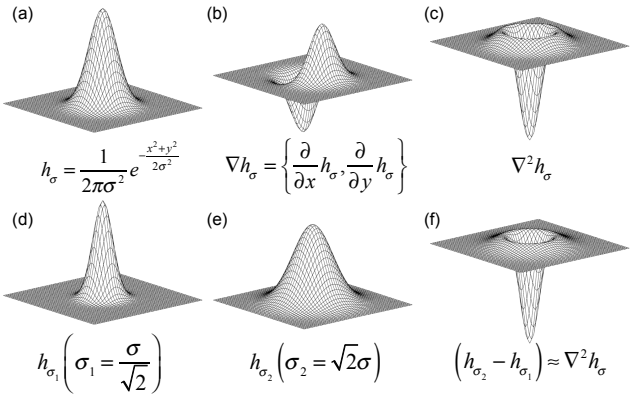

Figure 1. (a)–(c) Comparison of the Gaussian function h_σ, one of gradients ∇h_σ and Laplacian $\nabla^2 h_\sigma$. (d)–(f) two different Gaussians with σ_1 and σ_2 and a difference of the Gaussians.

shown in Figure 1, the Gaussian function (a), the *gradients* of the Gaussian (b) and the divergence of gradients, *Laplacian* (c) have been commonly used for filtering. Since convolutional filtering is a linear operation, the Laplacian of a Gaussian-filtered image $\nabla^2 [h_\sigma \otimes f]$ is identical to the image filtered with the Laplacian kernel $\left[\nabla^2 h_\sigma\right] \otimes f$. Downsampling the signals subsequent to the filtering builds an image pyramid.

In particular, the gradients $\nabla_{x,y}$ and the Laplacian ∇^2 have been commonly used to detect *edge* structures in images. The gradients are calculated from the first partial derivatives of x and y: $\nabla h_\sigma = \left\{ \frac{\partial}{\partial x} h_\sigma, \frac{\partial}{\partial y} h_\sigma \right\}$. At least two operators are required to detect the changes of local structures in images. See Figure 1. The gradients are *directional* edge detectors that require multiple operators for each direction; therefore, the gradient magnitude is used alternatively for detecting edges. Unlike the gradients, the Laplacian operator $\nabla \cdot \nabla h_\sigma = \nabla^2 h_\sigma = \left(\frac{\partial^2 h_\sigma}{\partial^2 x^2} + \frac{\partial^2 h_\sigma}{\partial^2 y^2} \right)$ is an *isotropic* edge detector which is invariant to rotation of the function in images. The Laplacian pyramid (Figure 2a) stores the bandpassed structural information of each frequency band. The pyramid has been used broadly for various edge-aware image processing [21].

Computing the Laplacians. Since convolution with large weighting functions is an expensive computation, the Laplacian of a Gaussian (LoG) can be approximated by simply taking a difference of two Gaussians (DoG) at different scales. See Figures 1(d)—(f). When the image scale of level $(l+1)$ reduces in a half scale of level l in the Gaussian pyramid, the DoG can approximate the LoG with high accuracy. Figures 1(c) and 1(f) compare the similarity of the LoG and the DoG. The backward computation of the gradients is notoriously expensive due to solving the Poisson equation, which may also introduce undesirable artifacts caused by non-integrable gradient fields. In contrast, the forward and backward computation of the Laplacian pyramid using the DoG is very efficient as they are virtually the operations of subtraction and summation, so the Laplacian pyramid has excellent potentials for patch-based image synthesis. In this paper, we employ the DoG-based approximation to compute the LoG efficiently. Note that for this reason, unlike other gradient-based methods such as Darabi et al. [11], our method requires no more additional computational cost, such as solving the Poisson equation, computed with traditional patch-based synthesis methods [29, 3].

3.2. Patch-Based Synthesis on a Laplacian Pyramid

The choice of the pyramid kernel operator is critical with respect to which information we derive from images. The Gaussian operator is effective in determining the base structures at each level of frequency. This operator is adopted in many image completion algorithms to achieve *spatial coherence* in searching correspondence and aggregating similarity, proposed by Wexler et al. [29] and Barnes et al. [3]. Contrary to the Gaussian, the gradient and the Laplacian operator are capable of searching *edge structures* of each level in images. The pyramid elements of derivatives are preserved to local regions in the spatial domain of the gradients or the Laplacian. The derivative image pyramids decompose edge localization at each level of frequency.

Recently, Darabi et al. [11] introduced a correspondence search that examines not only color but also *gradients* in the image pyramid. They presented that examining the first derivatives is clearly beneficial in searching correspondences of *structural coherence* in image completion. We were motivated to improve the seminal idea of leveraging derivatives with the Laplacian. Our edge-aware search of correspondence resembles the work of Darabi et al. with two main differences of *aggregated correspondence* and *rotation invariance* as follows.

Building a Laplacian Pyramid. The input of exemplar-based image completion is a color image I and a mask image M, which segregates an image into source region S and target region T. The goal of exemplar-based image completion is to complete the target region T of image I with contents from the source region S. To use a Laplacian pyramid in image inpainting, we first build a Gaussian pyramid G for image I following [21]:

$$\begin{aligned} G_0 &= I, \\ G_{i+1} &= \mathrm{downsample}(G_i) \quad (i < n), \end{aligned} \tag{1}$$

where G_i is the i-th scale in the Gaussian pyramid G, and the total number of scales is $n+1$, and $\mathrm{downsample}()$ is an operator that subsamples a filtered scale. The finest level of the Gaussian pyramid G_0 is the original image I. The i-th Gaussian scale G_{i+1} is a subsampled scale from the Gaussian-filtered one of the previous scale G_i.

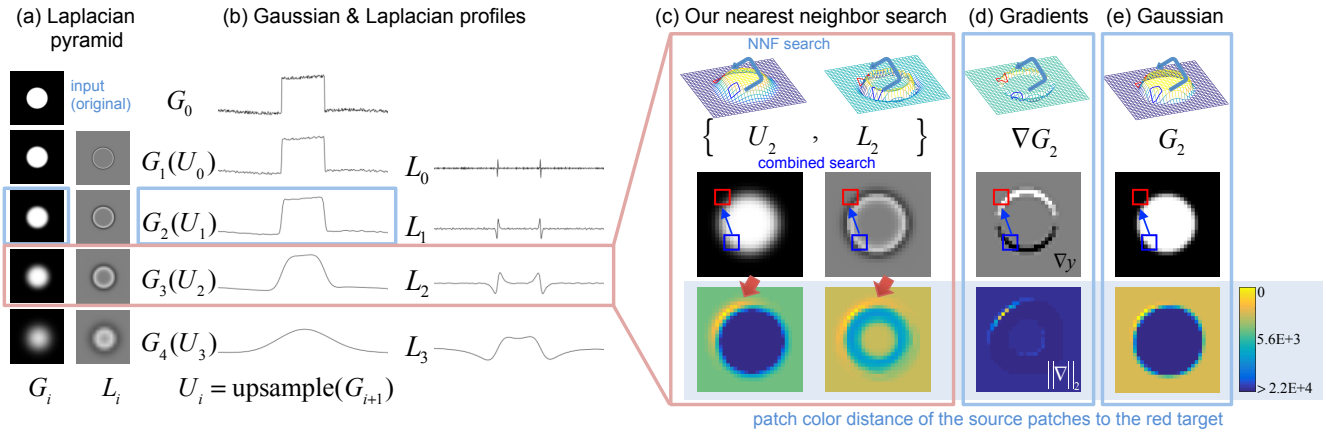

Figure 2. Comparison of correspondence search on the Gaussian, the gradients, and the Laplacian pyramid. (a) shows a Gaussian and a Laplacian pyramid of a circular function. (b) indicates 1D profiles of each level. (c), (d) and (e) show patch searches on level 2: the upsampled Gaussian U, the Laplacian L, gradients ∇G and the traditional Gaussian G, respectively. The top row of (c), (d), and (e) presents the 3D visualization of image structures. The middle row offers close-up images in the level. The bottom row visualizes an example of the source patch distances to the red target patch (4×4) over the rounded edge. Contrary to (d) and (e), our NNF search (c) compares the patch distances in U and L simultaneously, where the *aggregated correspondences* (indicated by red arrows) of both low-frequency base and high-frequency detail structures spread smoothly, giving assistance to random correspondence searches.

We then compute differences of the Gaussians to derive a Laplacian pyramid L:

$$\begin{aligned} U_i &= \text{upsample}(G_{i+1}), \\ L_i &= G_i - U_i \quad (i < n), \\ L_n &= G_n, \end{aligned} \tag{2}$$

where upsample() is an upsampling operator. To simplify our algorithm, we define an upsampled Gaussian pyramid U. The i-th Laplacian image L_i is the detailed structures between G_i and G_{i+1}. Since the number of levels in the Gaussian pyramid is finite, the coarsest level of the Laplacian pyramid L_n is the coarsest of the Gaussian pyramid G_n, called the residual, which corresponds to a tiny version of the image.

Our DoG-based Laplacian pyramid includes two image pyramids of the Gaussians and the Laplacians, shown in Figures 2(a) and (b). The frequency of the *base structure* is decomposed as intensity at each level in the Gaussian pyramid, while the frequency of the *edge structure* is localized as the Laplacians at each level of frequency in the Laplacian pyramid. Figure 2 offers an overview of our edge-aware correspondence search on the Laplacian pyramid, compared with a gradient-based search [11] (d) and a traditional Gaussian-based search (e) [29, 3]. We construct pyramids with a 5-by-5 blur kernel.

Aggregated Correspondence. Our main intuition on the selection of these two pyramids is the *aggregated correspondence* of the base and the edge structures. The bottom row in Figure 2(c) shows an example of the patch color distances of all the source patches to the target patch at the top-left round corner of a circular function at scales of U_2 and L_2. The area pointed by the red arrows indicates *aggregated* information of the base and the detailed structures in the both pyramids. Since the random search algorithm gradually steps forward to minimize the patch color distance in Equation (3), the aggregated edges of the base and the detail structures become significantly helpful clues for searching the correspondences of structural similarity due to the nature of the random search algorithm [11]. (The brighter colors in the figure indicate more similar patch color distances.) In contrast, the distance metrics on both the gradient magnitude $||\nabla||_2$ (d) and the Gaussian scale (e) weight only a few number of patches outstandingly high. Consequently most of the patches' structures are ignored by the weight when calculating the weighted sum of colors for vote (see Equation (4)). These two image pyramids of the gradients and the Gaussians are used by Darabi et al. [11] for the energy function of patch distance.

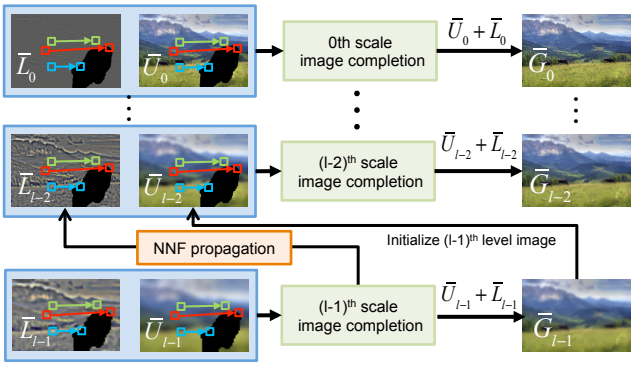

Figure 3. Overview of our implementation of Laplacian patch-based image synthesis.

Based on this observation, our objective strategy of

searching for correspondence not only minimizes the distances of low-frequency base structures but also preserves the distances of high-frequency detailed structures.

Our energy function resembles the iterative Expectation-Maximization (EM) algorithm following Wexler et al. [29] but with the main difference of using two image pyramids of the *upsampled* Gaussians and the Laplacians. In the expectation step, we find similar source patches (blue square in Figure 2c) for all target patches (red square). We use the random correspondence search algorithm, proposed by Barnes et al. [3], to approximate the nearest-neighbor fields of target patches with our distance metric as follows:

$$E_i(T,S) = \sum_{q \in T} \min_{p \in S} \left(\alpha D\left(U_{i,p}, U_{i,q}\right) + \beta D\left(L_{i,p}, L_{i,q}\right)\right), \tag{3}$$

where i is the current level, p and q are pixel locations of target region T and source region S respectively, $U_{i,p}$ and $L_{i,p}$ denote a patch of pyramid U and L at level i at position p respectively, D is the sum of square distances (SSD) between the CIELAB [9] colors of two patches, and finally α and β determine the ratio of low-frequency base scale U and high-frequency detailed scale L subject to $\alpha + \beta = 1$.

Rotation Invariance. As shown in Figure 2(d), image gradients $\nabla_x G$ and $\nabla_y G$ can detect only certain structural changes along the horizontal and the vertical direction since they are directional operators (Section 3.1). Unlike the gradient-based approach, we suggest using an *isotropic* edge operator of the Laplacian, which has the important advantage of being invariant to rotation [18]. The Laplacian responds equally to structure changes over edges in any direction in scales, detecting detailed structure more robustly. This isotropic characteristics of the Laplacians also avoid having to use multiple operators for the gradients to calculate the robust correspondence of local structures, allowing for computational efficiency.

Combined Vote. We update target patches of an upsampled Gaussian and a Laplacian image by blending nearest-neighbor source patches to maximize the similarity of target patches and source patches after searching correspondences. To accelerate the convergence, we perform weighted blending of scales. Patches of close color distances and patches close to the completion boundary are highly weighted. More details are provided in Section 4.

Laplacian Structure Reconstruction. Once the convergence of EM optimization is finished at level i, we are ready to propagate the current completion of $\overline{L}_i$ and $\overline{U}_i$ to the finer level $i-1$. Consequently these completions at level i can be used as the initial completion of level $i-1$. As shown in Figure 3 and Algorithm 1, the finer scales of the completed Gaussians $\overline{G}_i$ at level i can be obtained by summing the upsampled Gaussians $\overline{U}_i$ and Laplacians $\overline{L}_i$ at level i using Equation (2). The reconstructed Gaussians $\overline{G}_i$ is then upsampled to completed Gaussians $\overline{U}_{i-1}$ in consequence. However, the finer scales of the Laplacians $\overline{L}_{i-1}$ at level $i-1$ cannot be reconstructed by this manner. Since the Laplacian scale $\overline{L}_{i-1}$ is supposed to hold high-frequency details, we fill in the target region of the Laplacian scales $\overline{L}_{i-1}$ with information of the source regions at level $i-1$ by utilizing the NNF correspondence of the search patches $\overline{N}_i$ obtained at level i. In this way, we can complete the base and the detailed structures in these two image pyramid from the coarse to the finest level.

Algorithm 1 Laplacian-based image completion

Input: image I and mask image M
Output: result image $\overline{G}_0$

$G, U, L \leftarrow$ CONSTRUCTPYRAMID(I)
initialize $\overline{U}, \overline{L}$
for scale $i = n-1$ to 0 **do**
  for iteration $j = 0$ to m **do**
    $\overline{N}_i \leftarrow$ SEARCH$(\overline{U}_i, \overline{L}_i, U_i, L_i)$
    $\{\overline{U}_i, \overline{L}_i\} \leftarrow$ VOTE$(\overline{U}_i, \overline{L}_i, U_i, L_i, \overline{N}_i)$
  end for
  $\overline{G}_i \leftarrow \overline{U}_i + \overline{L}_i$
  if $i > 0$ **then**
    $\overline{U}_{i-1} \leftarrow$ UPSAMPLE$(\overline{G}_i)$
    $\overline{L}_{i-1} \leftarrow$ LAPLACIANRECONSTRUCTION$(\overline{N}_i, L_{i-1})$
  end if
end for

4. Implementation Detail

Vote. For the color voting stage, we inherit one from Wexler et al. [29]. Here we briefly describe our voting implementation to help readers understand the entire workflow. We first compute the similarity of a target patch at pixel q and its corresponding source patch at pixel p at level l as $\Psi(p,q,l) = e^{-\frac{D(p,q,l)}{2\sigma^2}}$, where σ determines the sensitivity of detecting similarity. In addition to similarity Ψ, we calculate confidence weights $\Lambda(q)$ at target pixel q that avoids boundary errors by assigning a higher confidence value to target points when they are closer to the completion boundary. This voting process assumes that the image inside the target region should be located outside the target region in the image. We combine these two metrics of similarity and confidence as a weight $w_q = \Psi(p,q,l)\Lambda(q)$ at target pixel q. For every target pixel q, we compute a weighted average c_q of the overlapping colors of $\tilde{q} \in Q$ from its NNFs $\overline{N}_{\tilde{q}}$ using weight $w_{\tilde{q}}$:

$$c_q = \frac{\sum_{\tilde{q} \in Q} w_{\tilde{q}} \overline{N}_{\tilde{q}}(q - \tilde{q})}{\sum_{\tilde{q} \in Q} w_{\tilde{q}}}, \tag{4}$$

where Q indicates the overlapping patches over the target pixel q.

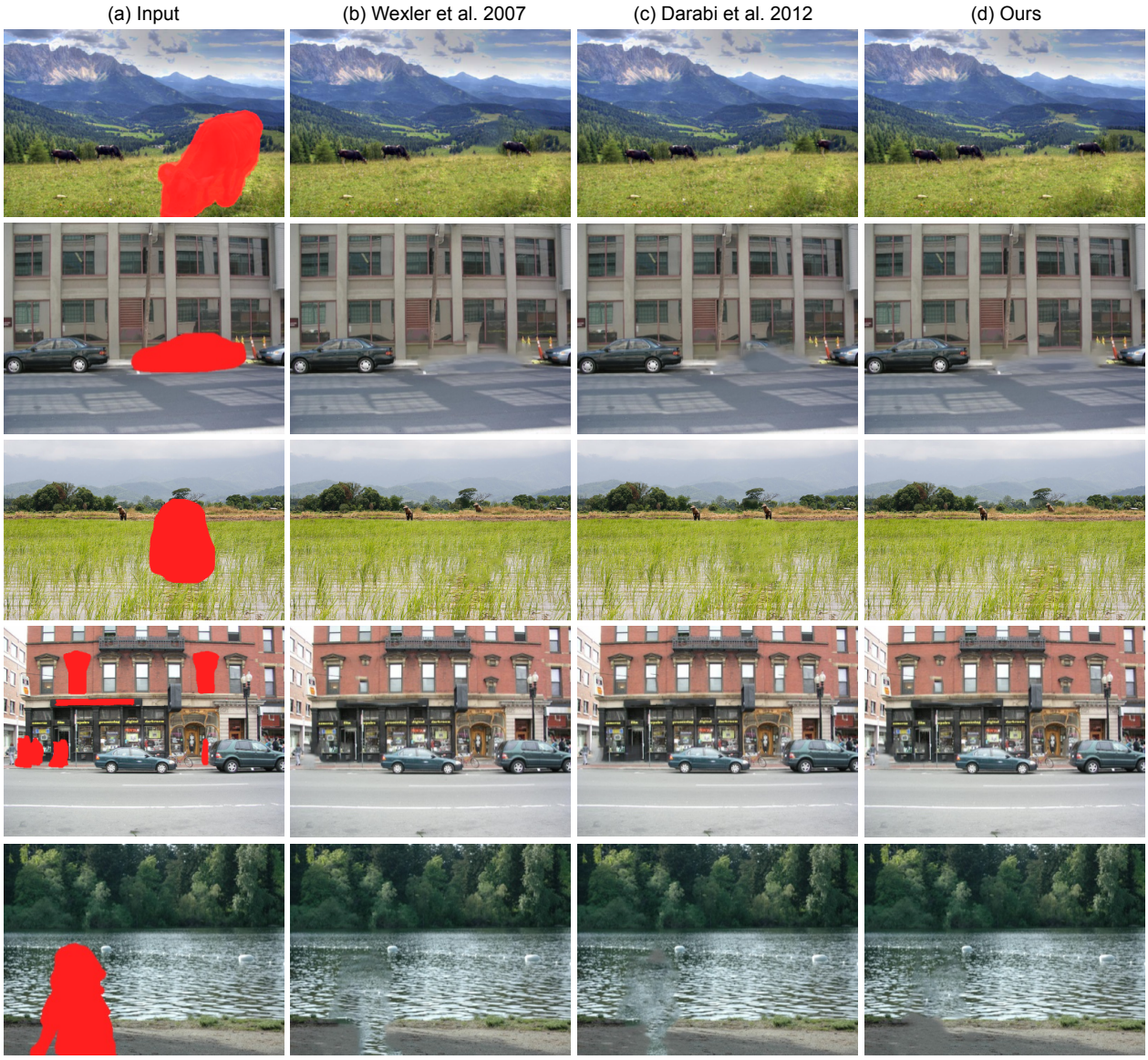

Figure 4. (a) Input images, where red region is to be completed, (b) Wexler et al. [29], (c) Darabi et al. [11], and (d) our method.

Figure 4	Wexler et al.	Darabi et al.	Ours
1^{st} row	105.35	367.65	112.40
2^{nd} row	56.00	314.54	59.27
3^{rd} row	127.45	732.96	138.31
4^{th} row	171.19	314.82	171.68
5^{th} row	76.67	314.17	76.44
Average	107.33	408.83	111.62

Table 1. Performance of three patch-based methods: Wexler, Darabi and ours with images shown in Figure 4 (unit: second).

5. Results

We implemented our Laplacian image synthesis in C++ on a machine with a 3.4GHz Intel i7-3770 CPU. Our unoptimized implementation runs in a genuinely single-threaded CPU-based manner. We compare our method with the current state-of-the-art inpainting methods, including other Laplacian-based inpainting approaches. We also evaluate the influence of the parameters, such as the patch size and β in Equation (3), in addition to the upsampled/naïve Gaussian pyramid structure in searching correspondence.

Figure 4 compares our method to the state-of-the-art methods, Wexler et al. [29] and Darabi et al. [11]. Our inpainting method synthesizes coherent base structures more robustly than other methods. In particular, the third-row results clearly show that our Laplacian patch-based method outperforms the gradient-based approach in terms of structure and detail. Our method effectively reconstructs the base structure of low spatial frequencies, e.g. the textures of rices and water in Figure 4, as our method leverages the informa-

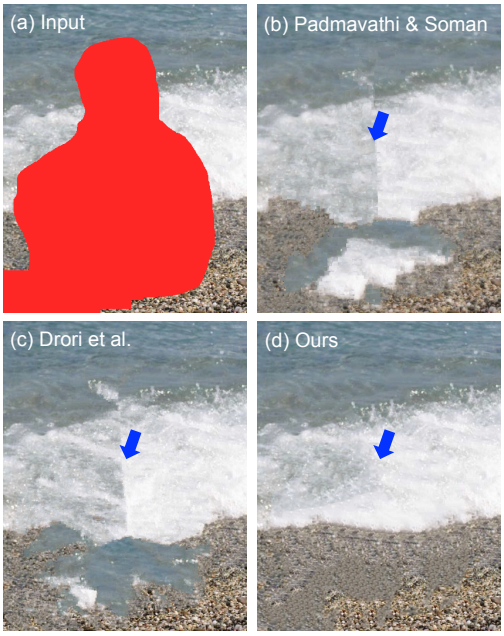

Figure 5. Comparison with other Laplacian-based methods. Both Padmavathi and Soman (b) and Drori et al. (c) contain seam artifacts due to collision of different local propagation while our method preserves global coherence in (d).

tion of the multiple frequency through the Laplacian pyramid.

Table 1 presents the performance of the three methods in computing results in Figure 4. We used our C++ implementation for Wexler et al. and the original code for Darabi et al.. Our method is implemented in C++. Wexler took 107.33 seconds per image in average, and Darabi took 408.83 seconds per image in average, while our method took 111.62 seconds per image in average. Our method is 3.66 times faster than the gradient-based method [11]. In detail, the step for Poisson reconstruction in Darabi took aver. 34.26 seconds per image, whereas our reconstruction step of the Laplacian pyramid took only aver. 0.02 seconds, revealing the computational efficiency of our method without sacrificing structural coherence in image synthesis.

Figure 5 shows the inpainted images of previous Laplacian-based approaches. While Padmavathi and Soman [20] (b) and Drori et al. [12] (c) utilize the Laplacian pyramid, they inherit the limitation of the heuristic search approach of Criminisi et al. In contrast, our method preserves global coherence (d) without suffering from the typical seam artifacts due to collision of different local propagation. See supplemental material for more results.

Varying Parameters. Figure 6 shows the consistency of our method under varying patch size. For this bungee example, the size of a patch is critical to detect and propagate a roof structure. The naïve Gaussian method with large patches misses to propagate the low-level edge structure (Figures 6(a), (b), and (c)). Thanks to the nature of the upsampled Gaussian pyramid, Figures 6(d), (e) and (f) implies that our method is able to consistently reconstructs low-edge structures with the diverse patch sizes.

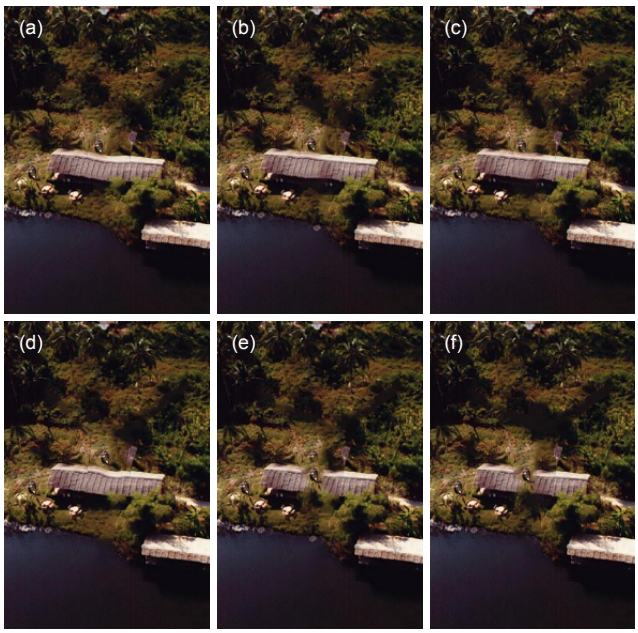

Figure 6. Bungee results of a Gaussian pyramid (d), (e) and (f) and a upscaled Gaussian pyramid (a), (b) and (c). As the patch size increases, the naïve Gaussian method fails to propagate low-level edge information. However, the upscaled Gaussian method maintain low-level structure as (b) shows.

Figure 7. The completed results with/without the Laplacian term. Our method with the non-zero Laplacian term preserves high frequency structures as shown in (b). However, without the Laplacian term, our method fails to keep high details in a region directed by blue arrows in (a).

We investigated the influence of the parameter β in Equation (3). See Figure 7. When β is zero, only the upsampled Gaussian term remains in Equation (3) and it results in over-smoothed results for high frequency textures as shown in (a). We found that β in the range between 0.1 and 0.5 produces plausible results.

Upsampled Gaussian vs. Gaussian. As shown in Figure 3, the summation of a Laplacian L_i and a upsampled Gaussian layer U_i in a Laplacian pyramid results in a Gaussian layer G_i, which is sharper than the upsampled

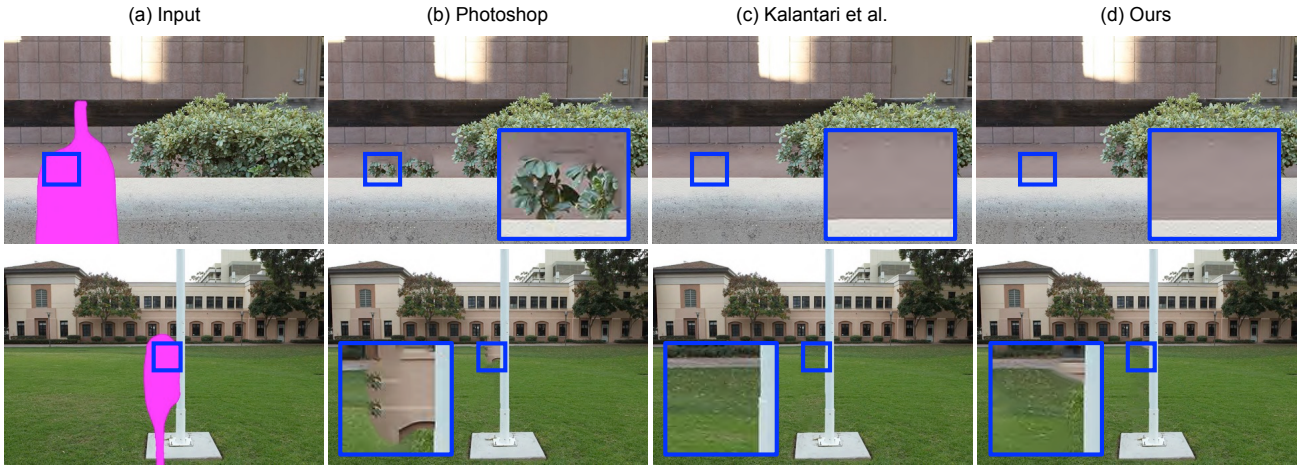

Figure 8. Comparison with Photoshop's contents-aware fill (b) and Kalantari et al. (c). Contents-aware fill often suffers from imperfect stitching, while our method maintains global coherence. However, our method inherits the limitation of single-layer patch-based synthesis, shown in the second row, while Kalantari's multi-layer approach outperforms ours with foreground and background patches.

Figure 9. We compare the Gaussian-domain search (b) and (e) and our upsampled Gaussian-domain search (c) and (f). We set parameters the same for both methods and the patch size is 9×9 for (b) and (c) and 5×5 for (e) and (f). Under the same patch size, (c) and (f) captures global structures, while (b) and (e) misses the image structure.

Gaussian U_i. When we compute the correspondence of the base and the edge structure, we have a domain option to use either U_i and G_i for searching the base structure. We compare them to verify the improvement on the global coherence. Figure 9 shows the results of this comparison. The Gaussian-domain search fails to capture edge structure in regions directed by blue arrows in (b) and (e) because of small patch size. Conversely, as shown in (c) and (f) our upsampled Gaussian-domain search successfully reconstructs global structures even using the same patch size. It demonstrates that adopting the upsampled Gaussian pyramid enables us to better preserve the low-level edge structure than using the naïve Gaussian pyramid does.

6. Conclusion and Future Work

We have presented an edge-aware patch-based image synthesis method based on the Laplacian pyramid. While the proposed method overcomes the shortcomings of the gradient-based synthesis such as directionality and heavy computational burden, the proposed method takes the advantages of the Laplacian pyramid properties such as aggregated correspondence and isotropic feature detection in order to improve searching correspondence with enhanced awareness of edge structures. To validate this proposed method, we demonstrate the effectiveness of this Laplacian-based approach over the state-of-the-art techniques with variety of experimental results.

Since our method stems from the single-layer patch synthesis approach [28], our approach is incapable of identifing foreground and background structure in the synthesized image. Compared to Photoshop's contents-aware fill, our results preserve global structure as coherent as possible shown in Figure 8(d). Yet, our method could often suffer from patch incoherence in the mixed situation of the foreground and background objects. Recently, Kalantari et al. [15] attempted to solve this problem by masking incoherent parts of patches, as shown in Figure 8(c).

Our method is based on a Laplacian pyramid. Thus, our method inevitably inherits natural drawbacks of a Laplacian pyramid. One of the drawbacks is that a Laplacian pyramid is not scalable compared to gradients. It makes our search space restricted into two degrees of freedom: rotations and translations without scales.

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.

References

[1] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, 1984.

[2] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. Image Processing (TIP)*, 10(8):1200–1211, 2001.

[3] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph. (TOG)*, 28(3):24:1–11, 2009.

[4] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein. The generalized Patchmatch correspondence algorithm. In *Proc. European Conf. Comput. Vision (ECCV)*, pages 29–43, 2010.

[5] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, 1975.

[6] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proc. ACM. SIGGRAPH*, pages 417–424, 2000.

[7] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. *Proc. IEEE. Conf. Comput. Vision and Pattern Recognition (CVPR)*, 2:1809–1824, 2005.

[8] P. Buyssens, M. Daisy, D. Tschumperlé, and O. Lézoray. Exemplar-based inpainting: Technical review and new heuristics for better geometric reconstructions. *IEEE Trans. Image Processing (TIP)*, 24(6), 2015.

[9] CIE. Colorimetry. CIE Pub. 15.2, Commission Internationale de l'Eclairage (CIE), Vienna, 1986.

[10] A. Criminisi, P. Perez, and K. Toyama. Object removal by eexemplar-based inpainting. In *Proc. IEEE. Conf. Comput. Vision and Pattern Recognition (CVPR)*, volume 2, pages 721–728, 2003.

[11] S. Darabi, E. Shechtman, C. Barnes, D. B. Goldman, and P. Sen. Image melding: Combining inconsistent images using patch-based synthesis. *ACM Trans. Graph. (TOG)*, 31(4):82:1–10, 2012.

[12] I. Drori, D. Cohen-Or, and H. Yeshurun. Fragment-based image completion. *ACM Trans. Graph. (TOG)*, 22(3):303–312, 2003.

[13] K. He and J. Sun. Image completion approaches using the statistics of similar patches. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 36(12):2423–2435, 2014.

[14] J.-B. Huang, S. B. Kang, N. Ahuja, and J. Kopf. Image completion using planar structure guidance. *ACM Trans. Graph. (TOG)*, 33(4):129:1–10, 2014.

[15] N. K. Kalantari, E. Shechtman, S. Darabi, D. B. Goldman, and P. Sen. Improving patch-based synthesis by learning patch masks. In *Proc. Int. Conf. Comput. Photography (ICCP)*, pages 1–8, 2014.

[16] N. Komodakis. Image completion using global optimization. In *Proc. IEEE. Conf. Comput. Vision and Pattern Recognition (CVPR)*, volume 1, pages 442–452, 2006.

[17] A. Levin, A. Zomet, and Y. Weiss. Learning how to inpaint from global image statistics. In *Proc. Int. Conf. Comput. Vision (ICCV)*, volume 1, pages 305–312, 2003.

[18] D. Marr and E. Hildreth. Theory of edge detection. *Proc. R. Soc. Lond.*, 207(1167):187–217, 1980.

[19] O. L. Meur, M. Ebdelli, and C. Guillemot. Hierarchical super-resolution-based inpainting. *IEEE Trans. Image Processing (TIP)*, 22(10):3779–3790, 2013.

[20] S. Padmavathi and K. P. Soman. Laplacian pyramid based hierarchical image inpainting. *Adv. Image and Video Processing (AIVP)*, 2(1):9–22, 2014.

[21] S. Paris, S. W. Hasinoff, and J. Kautz. Local laplacian filters: Edge-aware image processing with a laplacian pyramid. *ACM Trans. Graph. (TOG)*, 30(4):68:1–12, 2011.

[22] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 12(7):629–639, 1990.

[23] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992.

[24] J. Shen and T. F. Chan. Mathematical models for local nontexture inpaintings. *SIAM J. Appl. Math.*, 62(3):1019–1043, 2002.

[25] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum. Image completion with structure propagation. *ACM Trans. Graph. (TOG)*, 24(3):861–868, 2005.

[26] F. Voci, S. Eiho, N. Sugimoto, and H. Sekibuchi. Estimating the gradient in the Perona-Malik equation. *Trans. IEEE. Signal Processing Magazine*, 21(3):39–65, 2004.

[27] J. Weickert. Coherence-enhancing diffusion filtering. *Int. J. Comput. Vision (IJCV)*, 31(2):111–127, 1999.

[28] Y. Wexler, E. Shechtman, and M. Irani. Space-time video completion. In *Proc. IEEE. Conf. Comput. Vision and Pattern Recognition (CVPR)*, volume 1, pages 120–127, 2004.

[29] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)*, 29(3):463–476, 2007.

[30] Z. Xu and J. Sun. Image inpainting by patch propagation using patch sparsity. *IEEE Trans. Image Processing (TIP)*, 19(5):1153–1165, 2010.

[31] P. N. Yianilos. Data structures and algorithms for nearest neighbor search in general metric spaces. In *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 311–321, 1993.