Simultaneous Inpainting and Super-resolution Using Self-learning

Milind G. Padalkar
milind_padalkar@daiict.ac.in
Manjunath V. Joshi
mv_joshi@daiict.ac.in
Nilay Khatri
nilay_space@yahoo.co.in

Dhirubhai Ambani Institute of Information & Communication Technology (DA-IICT), Gandhinagar, Gujarat, India – 382007

Abstract

Past two decades have seen significant advancement in the techniques for scene completion and image super-resolution. Although many of the approaches solve these two problems by searching and processing of similar patches for estimating the unknown pixel values, the two problems have been addressed independently. In applications like creating immersive walkthrough systems or digital reconstruction of invaluable artwork, both inpainting and super-resolution of the given images are the preliminary steps in order to provide better visual experience. The usual practice is to solve these problems independently in a pipelined manner. In this paper we propose a unified framework to perform simultaneous inpainting and super-resolution. We construct dictionaries of image-representative low and high resolution patch pairs from the known regions in the test image and its coarser resolution. Inpainting of the missing pixels is performed using exemplars found by comparing patch details at a finer resolution, where self-learning is used to obtain the finer resolution patches by making use of the constructed dictionaries. The obtained finer resolution patches represent the super-resolved patches in the missing regions. Advantage of our approach when compared to other exemplar based inpainting techniques are (1) additional constraint in the form of finer resolution matching results in better inpainting and (2) inpainting is obtained not only in the given spatial resolution but also at higher resolution leading to super-resolution inpainting. Experiments on natural images show efficacy of the proposed method in comparison to state-of-the-art methods.

1 Introduction

Need for seamless removal of objects from images has motivated a number of works on image inpainting. The process of inpainting fills up the pixels in a region of interest in such a way that, in the context of the image, the filled region looks visually plausible. The missing pixels are filled either by gradually propagating information from outside the boundary of the region of interest or by making use of cues from similar patches. Based on the filling strategy the existing inpainting methods can be categorized into two important groups viz. methods based on solving partial differential equations (PDEs) [2, 13] and those based on exemplars [5, 6, 19, 22]. Among these, the exemplar based methods are more popular as processing of similar patches well synthesizes the texture inside the missing regions.

Pages 105.1-105.12
DOI: https://dx.doi.org/10.5244/C.29.105

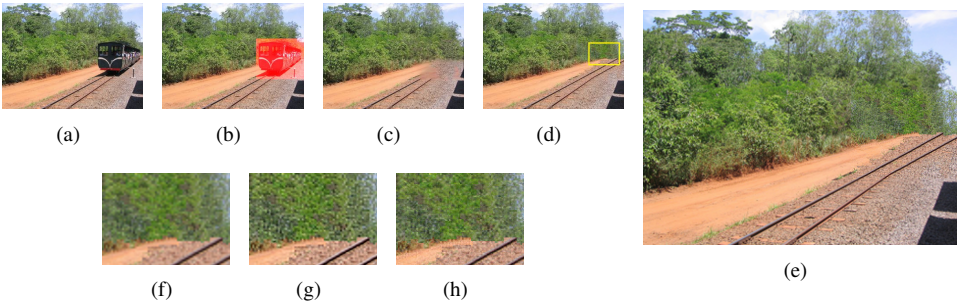


Figure 1: Simultaneous inpainting and super-resolution: (a) input; (b) region to be inpainted; (c) inpainting using planar structure guidance [13]; (d) inpainting using proposed method showing yellow box inside the inpainted region; (e) simultaneously inpainted and super-resolved image (by a factor of 2) using the proposed method with known regions upsampled using bicubic interpolation; (f)–(h) expanded versions after upsampling (the region marked by the yellow box in (d)) using various approaches viz. (f) bicubic interpolation, (g) Glasner *et al.*'s method [10] and (h) proposed method for super-resolution

While inpainting methods fill the missing pixels in the given image, the super-resolution (SR) methods obtain an upsampled version that preserves the high frequency details. In other words, the super-resolved image resembles the true image captured using a high-resolution (HR) camera. The techniques that estimate the HR image using multiple low-resolution (LR) images of same scene fall under the classical multi-image SR category [8, 21]. Example based SR is another category in which the correspondence between LR-HR patches is learnt from a database containing pairs of LR-HR images [9] or from the given image itself [10].

Therefore, one can see that by searching the similar patches we can estimate values of the missing pixels as well as perform resolution enhancement. In creating an immersive walkthrough system or digital reconstruction of invaluable artwork, the preliminary steps are to inpaint any existing cracks or damaged regions in the captured image and obtain high resolution details. This gives the viewers an enhanced visual experience. Similarly, in investigations based on photographs, inpainting can help in recreating a deliberately tampered region of the photograph while visual details could be enhanced using SR. However, when both inpainting and SR are to be performed on the given image, the usual practice is to first inpaint the missing regions and then independently super-resolve the inpainted image. One such example is the approach in [17]. Unlike our method it performs inpainting at a coarser resolution followed by independent SR to get the inpainted image at the original resolution.

In this paper we propose a unified method for image inpainting and SR. It is interesting to note that proposed inpainting method finds better exemplars at the original resolution and in the process also leads to SR of the inpainted region. In other words, we obtain SR as a consequence of inpainting, thus reducing the number of computations as compared to performing these operations independently. Note that our method does not use any kind of regularization as used by most of the SR approaches [20, 23]. We inpaint the missing pixels using exemplars found by comparing patch details at a finer (higher) resolution. Dictionaries of corresponding LR-HR patch pairs from the known region (i.e. region outside the missing pixels to be inpainted) are constructed and used in the compressive sensing framework to self-learn the HR of the patches that do not find a match in the dictionary or have missing pixels. The inpainting of patches in the original resolution uses an LR-HR relationship that

Symbols	Meaning
I_0, I_{-1}	Input image and its coarser resolution.
Ω_0, Ω_{-1}	Region to be inpainted in the input image I_0 and corresponding region in I_{-1}.
y_p	Patch of size $m \times m$ around a pixel $p \in I_0$.
y_p^u	Unknown missing pixels in the patch y_p that are to be inpainted i.e. $y_p^u \in \Omega_0$.
y_p^k	Known pixels in the patch y_p i.e. $y_p^k \in I_0 - \Omega_0$.
K	Number of candidate exemplars.
$y_{q_1}, \ldots, y_{q_K}$	Candidate exemplars corresponding to the patch y_p.
N	Number of patch pairs used for constructing the LR-HR dictionaries.
D_{LR}	Dictionary of low-resolution patches. Dimension: $m^2 \times N$.
D_{HR}	Dictionary of high-resolution patches. Dimension: $4m^2 \times N$.
$D_{LR_p}^k$	Dictionary of low-resolution patches containing only those rows that correspond to the known pixels y_p^k. Dimension: $\lvert y_p^k \rvert \times N$.
α	Sparse vector of size $N \times 1$.
Y_p	HR patch of size $2m \times 2m$ corresponding to LR patch y_p.
Y_p^u, Y_p^k	HR pixels in patch Y_p that correspond to the pixels y_p^u and y_p^k, respectively, in the LR patch y_p.
$Y_{q_1}, \ldots, Y_{q_K}$	HR patches corresponding to the candidate exemplars $y_{q_1}, \ldots, y_{q_K}$.
Y_q	Best match for Y_p among $Y_{q_1}, \ldots, Y_{q_K}$.
H_p	Final inpainted HR patch corresponding to the LR patch y_p.
L_p	Inpainted version of the LR patch y_p.

Table 1: Notation and description of the symbols used in this paper

1:	Construct LR-HR pair dictionaries using the known regions in I_0 and I_{-1}.
2:	Select highest priority patch $y_p = y_p^k \cup y_p^u$ for inpainting using method in [8]. Here, $y_p^k \in I_0$ and $y_p^u \in \Omega_0$.
3:	Search for candidate sources (exemplars) $y_{q_1}, \ldots, y_{q_K}$ in I_0.
4:	Self-learn HR patches $Y_{q_1}, \ldots, Y_{q_K}$ and Y_p using the constructed dictionaries: (a) Obtain $Y_{q_1}, \ldots, Y_{q_K}$ corresponding to $y_{q_1}, \ldots, y_{q_K}$. (b) Estimate Y_p corresponding to y_p.
5:	Find best exemplar Y_q in HR by comparing Y_p with $Y_{q_1}, \ldots, Y_{q_K}$.
6:	Obtain final inpainted HR patch H_p using Y_p and Y_q.
7:	Obtain inpainted LR patch L_p from H_p using transformation estimated from the constructed dictionaries and update Ω_0.
8:	Repeat steps 2–7 till all patches in Ω_0 are inpainted.

Table 2: Summary of the proposed approach

is also learnt from the constructed dictionaries, while the corresponding HR patches serve as simultaneously super-resolved patches. Thus, we obtain the inpainted region at the original resolution along with its super-resolved version as shown in figure 1. Note that our approach plausibly performs SR without introducing any blur or artefacts indicating better inpainting at the given resolution. We once again emphasize that the primary goal here is to obtain a well inpainted image. The super-resolved version is obtained as a by-product in the process of using an additional constraint that helps in finding a better source for inpainting.

2 Proposed approach

Natural images usually contain many self-similar patches. This cue has been used effectively by exemplar based inpainting methods, where search is done for the region to be filled up. However, when similar patches are unavailable, the inpainting may not be seamless resulting in graphical garbage. Even when similar patches are available, the best match may not always be a good source for inpainting. The reason is that the patch to be filled up has too

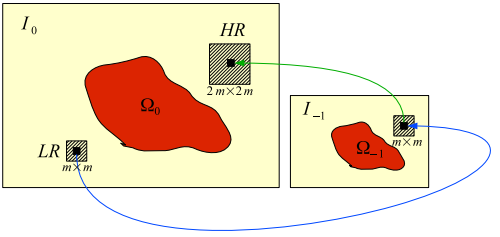


Figure 2: Finding LR-HR patch pairs using the given image I_0 and its coarser resolution I_{-1}.

little number of known pixels to obtain a reliable match. One may increase the patch size to have more number of known pixels. However, we may not find good matches for larger patches due to which the inpainted regions look implausible.

In exemplar based approaches, patch matching is done by discarding the missing pixels. Due to this it may happen that a better source for inpainting could be found among the patches other than the best matching patch. Therefore, it is desirable to consider the nearly best matching patches as candidate sources for inpainting without discarding them. Intuitively, by performing a detailed assessment of the patches to be filled, one can confidently determine which among the candidates is a better source for inpainting. In other words, if the HR details of the patches are made available, these can be used to find a reliable match which is a better exemplar. Khatri and Joshi [15] have shown that HR details can be self-learnt from the given image and its single coarser resolution. Drawing inspiration from [15], the proposed method estimates the HR details even for patches with missing pixels. Thus, the additional constraint of patch matching at original as well as finer resolution not only provides a better exemplar to fill the missing pixels but simultaneously also performs SR.

A brief description of the symbols used in this paper is given in table 1 and the proposed approach is summarized in table 2. The proposed approach starts with a given image I_0 having a region Ω_0 to be inpainted. We obtain the coarser resolution image I_{-1} by blurring and downsampling I_0 as done in [10]. Let Ω_{-1} denote the missing region in I_{-1}, which corresponds to Ω_0 as depicted in figure 2. For every $m \times m$ sized patch on the boundary of Ω_0, a data term and a confidence term denoting the presence of structure and proportion of known pixels, respectively, are calculated using the method proposed in [6]. The patch y_p around a pixel p for which the product of the data and confidence terms is the highest is then selected as the highest priority patch for inpainting. Let y_p^k and y_p^u denote the known and the unknown pixels in y_p, respectively. The patch y_p is then compared with every $m \times m$ sized patch in the known region $I_0 - \Omega_0$ using sum of squared difference (SSD) by considering only the pixels corresponding to y_p^k. We then obtain K best matches denoted as $y_{q_1}, \ldots, y_{q_K}$ representing the candidates. The exemplar based methods use $K = 1$ to obtain the best match, whereas our method considers more candidate matches by setting $K > 1$ in order to find a better exemplar. These patches are then used in obtaining HR patches.

Consider an LR patch of size $m \times m$ in the known region $I_0 - \Omega_0$. We can obtain the corresponding $2m \times 2m$ sized HR patch in the same resolution by considering the coarser resolution I_{-1} as illustrated in figure 2. Although not all LR patches can find a good match in the coarser resolution, we use this methodology to create dictionaries of image-representative LR-HR patch pairs, with the help of which a good match is estimated for any LR patch in the known region. We also learn the HR of an LR patch y_p with missing pixels (i.e. $y_p^u \in \Omega_0$) by making use of these LR-HR patch pairs. Simultaneous inpainting and SR of the missing

pixels in then performed by refining the estimated HR of y_p using HR of the best candidate among $y_{q_1}, \ldots, y_{q_K}$ and an LR-HR relationship learnt from the known region. Thus, we make use of self-learning while obtaining the HR patches of inpainting region which are then used to obtain the corresponding inpainted LR patches. In what follows we provide the details of (a) constructing image-representative LR-HR patch pair dictionaries, (b) self-learning the HR patches and (c) simultaneous inpainting and SR of missing pixels.

2.1 Constructing image-representative LR-HR patch pair dictionaries

To obtain the image-representative LR-HR patch pairs, we consider every $m \times m$ sized patch in the known region $I_0 - \Omega_0$. For each of these patches we find the best match by searching for similar patches in $I_{-1} - \Omega_{-1}$ (see figure 2). We then get the corresponding HR in $I_0 - \Omega_0$. Here, every LR patch will be mapped to exactly one HR patch. However, an HR patch may be mapped by many LR patches (when the LR patches are similar).

We then plot the histogram of mapped HR patches versus the frequency of mapping to determine of the most mapped HR patches. The HR patches that are highly mapped indicate repetitiveness of the LR patches and are therefore appropriate for representing the image patches. On the other hand, the HR patches having less frequency of mapping are less likely to represent the patches inside the region to be filled up. Such patches are therefore discarded. The highly mapped HR patches form the HR dictionary of size $4m^2 \times N$ and the corresponding $m \times m$ sized patches in $I_{-1} - \Omega_{-1}$ form the LR dictionary of size $m^2 \times N$. Here N is the number of highly mapped patches such that $N >> 4m^2$. Note that this pair of dictionaries do not have LR-HR pairs for every patch in the known region of I_0.

2.2 Self-learning HR patch

For an LR patch whose match is directly available in the LR dictionary, the corresponding patch in the HR dictionary is the required HR patch. For other LR patches we estimate a good match using a linear combination of few patches in the LR dictionary. When a signal is known to be sparse, the compressive sensing (CS) theory [4] provides a method to obtain the sparse representation. In our case, an LR patch y whose HR version needs to be estimated, can be sparsely represented using the LR dictionary D_{LR} such that:

$$y = D_{LR} * \alpha, \tag{1}$$

where α is a sparse vector of size $N \times 1$ and y represents the lexicographically ordered patch of size $m^2 \times 1$. In CS framework, the sparse vector is obtained by posing the problem as:

$$\min ||\alpha||_{l_1}, \quad \text{subject to} \quad y = D_{LR} * \alpha, \tag{2}$$

where $||\alpha||_{l_1}$ corresponds to $\sum_{j=1}^{N} |\alpha_j|^1$ which is minimized using standard optimization tools [3]. In this way, we obtain good matches from the already available LR dictionary itself. Assuming the LR-HR patch pairs to have the same sparseness and using the estimated sparse coefficients (α), the corresponding HR patch Y of size $4m^2 \times 1$ is obtained as follows:

$$Y = D_{HR} * \alpha, \tag{3}$$

where D_{HR} denotes the HR dictionary. The pixels in Y_p are rearranged to get a patch of size $2m \times 2m$ by reversing the operation that was used to obtain the lexicographical ordering.

2.3 Simultaneous inpainting and SR of missing pixels

With the knowledge of LR-HR patch pair dictionaries and self-learning HR patches for LR patches that do not find a good match, we now describe how the missing pixels are inpainted.

As already discussed, for the patch y_p selected for inpainting having known pixels y_p^k and missing pixels y_p^u, we have the K best matches $y_{q_1},\ldots,y_{q_K}$ that are candidate sources for inpainting. For each of these candidates, the corresponding HR patches viz. $Y_{q_1},\ldots,Y_{q_K}$ are self-learnt using the method described in section 2.2 by replacing $y = y_{q_i}$, $\alpha = \alpha_{q_i}$ and $Y = Y_{q_i}$, $i = 1,\ldots,K$. Note that these candidate LR patches have no missing pixels and, therefore, the estimated HR patches represent the true HR versions of the patches $y_{q_1},\ldots,y_{q_K}$.

The patch y_p that needs to be inpainted has missing pixels y_p^u. Therefore, one cannot directly obtain the corresponding HR patch. However, the known pixels y_p^k can be represented using a reduced LR dictionary $D_{LR_p}^k$ which consists of only those rows in D_{LR} that correspond to the pixels y_p^k depending on which of the pixels in y_p are missing. Here $D_{LR_p}^k$ is of size $|y_p^k| \times N$ where $|y_p^k|$ denotes the number of known pixels in y_p. Again, the method described in section 2.2 is used to obtain the HR patch Y_p corresponding to y_p, by replacing $y = y_p^k$, $\alpha = \alpha_p$, $D_{LR} = D_{LR_p}^k$ and $Y = Y_p$. Note that in order to obtain Y_p we use the complete HR dictionary D_{HR} of size $4m^2 \times N$ and hence Y_p has the size of $2m \times 2m$, i.e. it has no missing pixels. Since Y_p is obtained by considering only the known pixels $y_p^k \in y_p$ and the corresponding dictionary D_{LR}^k, the pixels Y_p^k that correspond to y_p^k represent true HR pixels. Likewise, the HR pixels Y_p^u that correspond to y_p^u provide a better approximation to the missing HR pixels due to the use of many similar and representative patches.

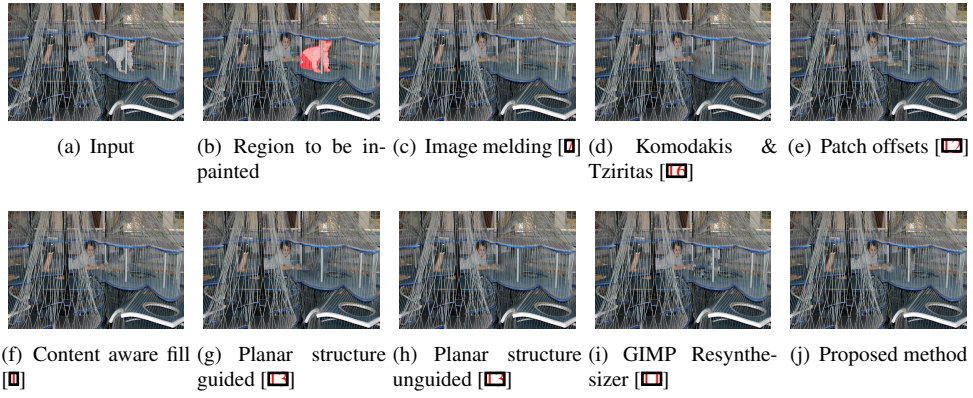
The final HR patch selection for missing regions is done as follows. We compare each of the HR patches $Y_{q_1},\ldots,Y_{q_K}$ with Y_p and choose the one having minimum SSD as Y_q. As the pixels in Y_p^u represent approximate but not true HR version of the missing pixels we replace them with those in Y_q in which all pixels represent true HR. The resulting patch H_p is final HR patch which is then used to obtain L_p representing the inpainted version of the patch y_p.

In order to obtain L_p from H_p we need the HR to LR transformation. In our case, blurring and downsampling is used to obtain coarser resolution I_{-1} from I_0 as done in [10]. Hence the same operation is used to obtain L_p from H_p. However, if the point spread function (PSF) of the camera is available, one can use it and perform downsampling to obtain the coarser resolution patches. Alternatively, if one uses I_{-1} that is captured using a camera, then the HR to LR transformation can be estimated from the available dictionaries having true LR-HR patch pairs to get L_p from H_p. Once the LR-HR dictionary pair is available we can model each LR pixel lr_i as a linear combination of 4 HR pixels $hr_i^{00}, hr_i^{01}, hr_i^{10}$, and hr_i^{11} as follows:

$$lr_i = [hr_i^{00}\ hr_i^{01}\ hr_i^{10}\ hr_i^{11}][a_{00}\ a_{01}\ a_{10}\ a_{11}]^T, \tag{4}$$

where a_{00}, a_{01}, a_{10} and a_{11} are the coefficients of the linear combination. Using the pixels in the LR-HR pair dictionaries in equation (4) these coefficients can be estimated in the least-squares sense. We can then obtain L_p from H_p by making use of the estimated coefficients.

We now have both LR and corresponding HR patches which are inpainted. The patch H_p is now placed appropriately in the upsampled image to obtain SR of the inpainted region. This process is repeated to inpaint the entire missing region Ω_0. Note that in every iteration only the missing pixels y_p^u in the selected patch y_p are inpainted and the missing region Ω_0 is updated accordingly. The order in which the patches are selected for filling is based on presence of structure and number of known pixels. This helps in propagating the structure inside the missing regions as a result of which the global structure is preserved. One may



(a) Input (b) Region to be inpainted (c) Image melding [7] (d) Komodakis & Tziritas [16] (e) Patch offsets [12]

(f) Content aware fill [1] (g) Planar structure guided [13] (h) Planar structure unguided [13] (i) GIMP Resynthesizer [11] (j) Proposed method

Figure 3: Results of inpainting the marked region corresponding to one of the kids in the cage

also super-resolve all the patches in the known region by self-learning the corresponding HR patches as explained in section 2.2. This will result in HR image where both known and inpainted regions are super-resolved.

3 Experimental results

In this section we present the results of experiments performed on the natural scene dataset available in [14]. The dataset also contains results of the state-of-the-art methods for image inpainting viz. image melding [7], Photoshop CS5 content aware fill [1], statistics of patch offsets [12], GIMP Resynthesizer plugin [11], planar structure guidance [13, 22] and the method by Komodakis and Tziritas [16]. We compare the results of our proposed method with these methods. The number of candidate matches considered in our implementation is $K = 5$ and the patch size is taken to be $m = 7$. The comparative results are presented in figures 3–8 which are discussed below.

Figure 3 shows the results of inpainting the marked region corresponding to one of the kids in the cage. The outline of the kid is visible and the bars show inconsistent bending in the inpainted results shown in figures 3(c)–3(d). An extra arm can be seen in figure 3(e) while some artefacts can be seen in figures 3(f) and 3(i). The results in figures 3(g)–3(h) are not only blurred, but also show inconsistency in the inpainted bars. The inpainted region in the proposed method shown in figure 3(j) looks visually better when compared to other approaches. The results of inpainting people and a vehicle in front of a shop are shown in figure 4. An implausible inpaintng of the region occluded by the vehicle can be seen in figures 4(c)– 4(d) and 4(f)–4(i). Similarly none of the results in figures 4(c)–4(i) show completion of the advertisement board occluded by the vehicle. Observe that the inpainting result of the proposed method displayed in figure 4(j) is not only plausible within the region but also well restores the advertisement board.

Figure 5 shows the inpainting of a table and chairs in a restaurant. In each of the inpainted results in figures 5(c)–4(g) a part of either chairs or table is visible, while the results in figures 5(h)–5(i) show improper inpainting of the brown tiles. The result of the proposed approach depicted in figure 5(j) does not show any artefacts of table or chairs in the inpainted regions

(a) Input

(b) Region to be inpainted

(c) Image melding [8]

(d) Komodakis & Tziritas [14]

(e) Patch offsets [12]

(f) Content aware fill [8]

(g) Planar structure guided [13]

(h) Planar structure unguided [13]

(i) GIMP Resynthesizer [11]

(j) Proposed method

Figure 4: Results of inpainting people and vehicle near the shop

(a) Input

(b) Region to be inpainted

(c) Image melding [8]

(d) Komodakis & Tziritas [14]

(e) Patch offsets [12]

(f) Content aware fill [8]

(g) Planar structure guided [13]

(h) Planar structure unguided [13]

(i) GIMP Resynthesizer [11]

(j) Proposed method

Figure 5: Results of inpainting the table and chairs in a restaurant

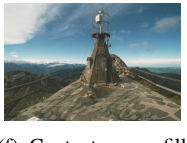
(a) Input

(b) Region to be inpainted

(c) Image melding [8]

(d) Komodakis & Tziritas [14]

(e) Patch offsets [12]



(f) Content aware fill [8]

(g) Planar structure guided [13]

(h) Planar structure unguided [13]

(i) GIMP Resynthesizer [11]

(j) Proposed method

Figure 6: Results of inpainting benches on the hill-top

(a) Input (b) Region to be inpainted (c) Image melding [9] (d) Komodakis & Tziritas [16] (e) Patch offsets [13]

(f) Content aware fill [8] (g) Planar structure guided [13] (h) Planar structure unguided [13] (i) GIMP Resynthesizer [11] (j) Proposed method

Figure 7: Results of inpainting people in front of the trucks

and the inpainted tile region looks acceptable. Another result in figure 6 shows the inpainting of benches on a hill-top. The result in figure 6(d) shows unrealistic criss-cross shadows of the fence, while those in figures 6(c), 6(f) and 6(h) have shadow of the fence in the right-half of the image, which is undesirable. The result shown in figure 6(g) is clearly not consistent with the known regions. Similarly, figure 6(e) has the door extended downwards that unrealistically cuts through the floor, while figure 6(i) appears to have a visible seam on the boundary of the inpainted region. Note that the result of the proposed method in figure 6(j) does not have any unrealistic shadows and is seamlessly inpainted. The texture of the inpainted region matches well with the region surrounding it.

In order to show the effectiveness of our approach on inpainting the region with low contrast we now consider another example. These results are shown in figure 7. We see that in the result of the proposed method shown in figure 7(j), the bumper of the truck in the left side of the image is well inpainted. None of the results shown in figures 7(c)–7(h) show completion of the bumper region. Similarly, in figures 7(c)–7(i), the edge of the pavement below the bumper does not appear to be convincingly inpainted, whereas figure 7(j) looks better inpainted. From the all results shown in figures 3–7, it is clear that our method performs better when compared to state-of-the-art approaches. Hence one can say that adding an additional constraint of matching patches at higher resolution results in better inpainting.

In order to show the effectiveness of our approach in super-resolving in addition to inpainting, we also present a result showing SR in figure 8. The inpainted and super-resolved region is compared with Glasner *et al.*'s approach [10] where the SR is performed on our inpainted result at the original resolution. Note that SR approaches super-resolve only what is available i.e. regions having no missing pixels, whereas the missing pixels are estimated and also super-resolved in our approach. Hence, our approach not only inpaints but also

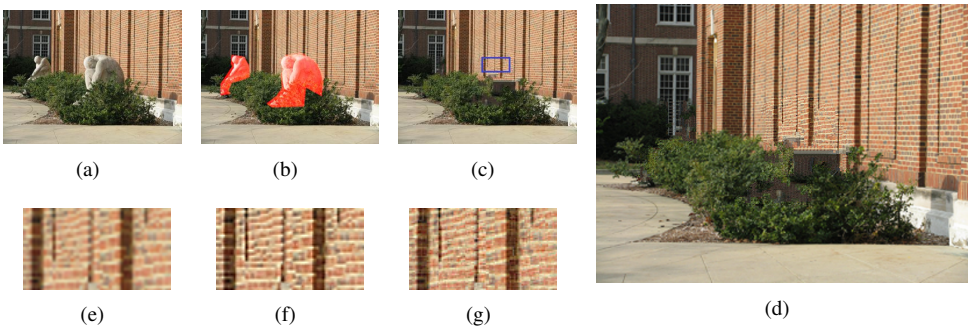


Figure 8: Result showing simultaneous inpainting and SR: (a) input; (b) regions to be inpainted ; (c) inpainting using proposed method showing blue box inside one of the inpainted regions; (d) simultaneously inpainted and super-resolved image (by a factor of 2) using the proposed method with known regions upsampled using bicubic interpolation; (e)–(g) expanded versions after upsampling (the region marked by the blue box in (c)) using various approaches viz. (e) bicubic interpolation, (f) Glasner *et al.*'s method [10] and (g) proposed method for super-resolution

reconstructs high resolution of the unknown region with missing pixels. We display the inpainted result in figure 8(c) and simultaneous SR in figure 8(d) obtained using the proposed method. The expanded version after upsampling one of the inpainted regions (shown by the blue box in figure 8(c)) using bicubic interpolation and Glasner *et al.*'s method [10] for SR are depicted in figures 8(e) and 8(f), respectively. Looking at the results, we see that the super-resolved region shown in figure 8(g) is comparable to the SR result shown in figure 8(f). Also, the simultaneously super-resolved region shows greater details than simply upsampling the inpainted region using bicubic interpolation as shown in figure 8(e).

4 Conclusion

We have presented a unified approach to perform simultaneous inpainting and SR. By using an additional constraint of matching patches at the original resolution as well as at the higher resolution, we not only obtain better source patches for inpainting but also have the corresponding super-resolved version. A comparison with the state-of-the-art inpainting methods shows that the inpainted results of the proposed method are indeed better. Also, the simultaneously super-resolved regions are comparable to the SR of the inpainted regions obtained using the method in [10] and also show greater details than those obtained by upsampling the inpainted regions using bicubic interpolation.

Acknowledgement

This work is a part of the project *Indian Digital Heritage (IDH) - Hampi* sponsored by Department of Science and Technology (DST), Govt. of India.
(Grant No: NRDMS/11/1586/2009/Phase-II)

References

[1] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. Patchmatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.*, 28(3):24:1–24:11, July 2009.

[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*, SIGGRAPH '00, pages 417–424, 2000.

[3] Emmanuel J. Candès and Justin Romberg. $l_1 - MAGIC$: Recovery of sparse signals via convex programming, 2005. URL http://users.ece.gatech.edu/~justin/l1magic/#links.

[4] Emmanuel J. Candès and Michael B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, March 2008.

[5] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Object removal by exemplar-based inpainting. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2:721, 2003.

[6] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Region filling and object removal by exemplar-based inpainting. *IEEE Transactions on Image Processing*, 13(9): 1200–1212, January 2004.

[7] Soheil Darabi, Eli Shechtman, Connelly Barnes, Dan B. Goldman, and Pradeep Sen. Image melding: Combining inconsistent images using patch-based synthesis. *ACM Trans. Graph.*, 31(4):82:1–82:10, July 2012.

[8] Sina Farsiu, M. Dirk Robinson, Michael Elad, and Peyman Milanfar. Fast and robust multiframe super resolution. *Image Processing, IEEE Transactions on*, 13(10):1327–1344, Oct 2004.

[9] William T. Freeman, Thouis R. Jones, and Egon C. Pasztor. Example-based super-resolution. *Computer Graphics and Applications, IEEE*, 22(2):56–65, Mar 2002.

[10] Daniel Glasner, Shai Bagon, and Michal Irani. Super-resolution from a single image. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 349–356, Sept 2009.

[11] Paul Francis Harrison. Gimp resynthesizer plugin, 2011. URL http://www.logarithmic.net/pfh/resynthesizer.

[12] Kaiming He and Jian Sun. Statistics of patch offsets for image completion. In *Proceedings of the 12th European Conference on Computer Vision - Volume Part II*, ECCV'12, pages 16–29, 2012. ISBN 978-3-642-33708-6.

[13] Jia-Bin Huang, Sing Bing Kang, Narendra Ahuja, and Johannes Kopf. Image completion using planar structure guidance. *ACM Trans. Graph.*, 33(4):129:1–129:10, July 2014.

[14] Jia-Bin Huang, Sing Bing Kang, Narendra Ahuja, and Johannes Kopf. Dataset for image completion using planar structure guidance, 2014. URL https://sites.google.com/site/jbhuang0604/publications/struct_completion.

[15] Nilay Khatri and Manjunath V. Joshi. Image super-resolution: Use of self-learning and Gabor prior. In *Proceedings of the 11th Asian Conference on Computer Vision - Volume Part III*, ACCV'12, pages 413–424, 2013.

[16] Nikos Komodakis and Georgios Tziritas. Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *Image Processing, IEEE Transactions on*, 16(11):2649–2661, Nov 2007.

[17] Olivier Le Meur and Christine Guillemot. Super-resolution-based inpainting. In *Computer Vision – ECCV 2012*, volume 7577 of *Lecture Notes in Computer Science*, pages 554–567. Springer Berlin Heidelberg, 2012.

[18] Simon Masnou and Jean-Michel Morel. Level lines based disocclusion. *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, pages 259–263, 1998.

[19] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03, pages 313–318, 2003.

[20] Pulak Purkait and Bhabatosh Chanda. Super resolution image reconstruction through bregman iteration using morphologic regularization. *Image Processing, IEEE Transactions on*, 21(9):4029 –4039, sept. 2012.

[21] Jing Tian and Kai-Kuang Ma. A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342, 2011.

[22] Yonatan Wexler, Eli Shechtman, and Michal Irani. Space-time completion of video. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(3):463–476, March 2007.

[23] Qiangqiang Yuan, Liangpei Zhang, and Huanfeng Shen. Regional spatially adaptive total variation super-resolution with spatial information filtering and clustering. *Image Processing, IEEE Transactions on*, 22(6):2327–2342, June 2013.