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

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 (SR). The main focus of this paper is inpainting, i.e. to remove objects in photographs and replace them with visually plausible backgrounds. 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.

The proposed approach starts with a given test 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 [2]. We then construct dictionaries of image-representative low and high resolution (LR-HR) patch pairs from the known regions in the test image I_0 and its coarser resolution I_{-1}. The inpainting of the missing pixels is performed using exemplars found by comparing patch details at a finer resolution. Here, self-learning [4] 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. The proposed approach is summarized in table 1 and one result is shown in figure 2.

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 [1]. 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 1: Summary of the proposed approach

Consider an LR patch of size $m \times m$ in the known region $I_0 - \Omega_0$. By searching for a similar $m \times m$ sized patch in the coarser resolution I_{-1} we can obtain the corresponding $2m \times 2m$ sized HR patch as illustrated in figure 1. Although not all LR patches can find a good match in the coarser resolution, we use this methodology to create dictionaries D_{LR} and D_{HR} of image-representative LR-HR patch pairs.

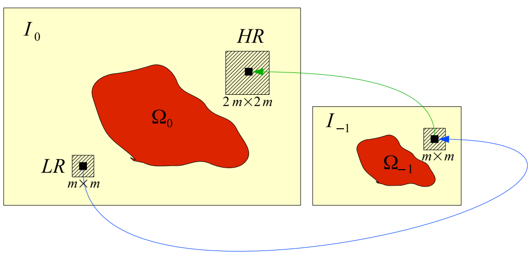

Figure 1: Finding LR-HR patch pairs using the given image I_0 and its coarser resolution I_{-1}.

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 in [1]. 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. The patch y_p is compared with every $m \times m$ sized patch in the known region $I_0 - \Omega_0$ to obtain K best matches denoted as $y_{q_1}, \ldots, y_{q_K}$ representing the candidate exemplars. The key idea is to compare the HR details of y_p with those of $y_{q_1}, \ldots, y_{q_K}$ to identify the best source for inpainting. Here, the HR patch Y corresponding to an LR patch y is self-learnt using the LR-HR patch pair dictionaries as follows:

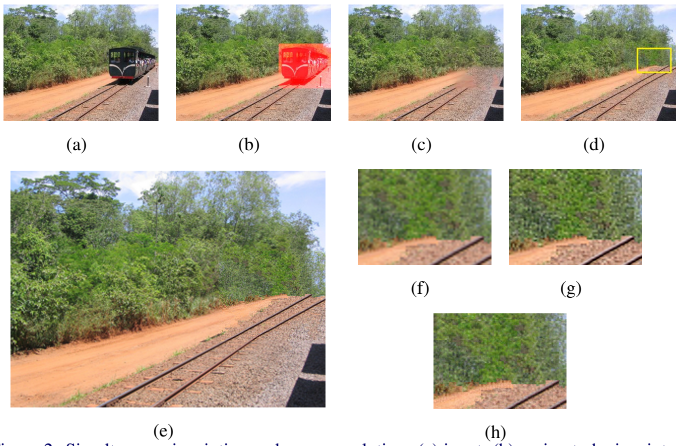

Figure 2: Simultaneous inpainting and super-resolution: (a) input; (b) region to be inpainted; (c) inpainting using planar structure guidance [3]; (d) inpainting using proposed method showing yellow box inside the inpainted region; (e) simultaneously inpainted and super-resolved image (by a factor of 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 [2] and (h) proposed method for super-resolution.

$$Y = D_{HR} * \alpha, \tag{1}$$

where, α is the sparse representation obtained by optimizing

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

For an LR patch y_p with missing pixels y_p^u and known pixels y_p^k, the corresponding sparse representation is obtained by using $D_{LR} = D_{LR_p}^k$ and $y_p = y_p^k$ in equation (2). Here $D_{LR_p}^k$ consists of only those rows in D_{LR} that correspond to the known pixels y_p^k in y_p. The complete HR dictionary D_{HR} and the obtained sparse representation are then used in equation (1) to get the HR details Y_p. Similarly HR $Y_{q_1}, \ldots, Y_{q_K}$ for candidates are obtained. We then compare each of the HR patches $Y_{q_1}, \ldots, Y_{q_K}$ with Y_p and choose the one having minimum sum of squared distance as Y_q. The pixels Y_p^u in the HR patch Y_p (that correspond to pixels y_p^u of the LR patch y_p) are now replaced by corresponding pixels in Y_q to obtain the final inpainted HR patch H_p. The corresponding LR patch L_p which is the inpainted version of y_p is obtained by using the HR to LR transformation, which in our case is blurring and downsampling or can be estimated from the constructed dictionaries. This process is repeated to inpaint the entire missing region Ω_0 such that in every iteration only the missing pixels y_p^u in the selected patch y_p are inpainted and Ω_0 is updated accordingly.

[1] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based inpainting. *Trans. Image Proc.*, 13(9): 1200–1212, Jan. 2004.

[2] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *ICCV 2009*, pages 349–356.

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

[4] N. Khatri and M. V. Joshi. Image super-resolution: Use of self-learning and Gabor prior. In *ACCV 2012*, pages 413–424.