Super Resolution from a Single Image based on Total Variation Regularization

Hiroki Tsurusaki*, Masashi Kameda and Prima Oky Dicky Ardiansyah

*Graduate School of Software and Information Science, Iwate Prefectural University
152-52, Sugo Takizawa Iwate, 020-0693, Japan

*Corresponding Author: g231l026@s.iwate-pu.ac.jp

Abstract

In this paper, we propose an approach to generate super-resolution from a single image using iterative total variation regularization. In the iterative process, the proposed method generates images representing skeletons and textures using pre-defined regulation parameters. We create a hybrid texture from the resulted textures and eliminate its non-edge regions to generate a high-resolution magnified image. Our experimental results show that the proposed approach generates magnified images with fine textures over the existing methods.

Keywords: Image Magnification, Super Resolution, Total Variation Regularization.

1. Introduction

Various methods have been proposed to generate super-resolution from a single image. They can be broadly classified into methods based on interpolation [1], example based super resolution [2][3] and single image super resolution [4][5][6]. While interpolation provides the simplest process, this method is unable to reconstruct the high frequency details. The example based super resolution makes low resolution and high resolution image patches in training database. A magnified image by this method depends on training database. And when the training data do not exist, we cannot use this method. Then we need single image super resolution, which not use training database.

Glasner’s method [4] may generate satisfied outputs if and only if there exist relatively similar images to the target images. Singh’s method [5] is edge reconstruction based on ramp transform [7] and regularized backprojection algorithm. The super-resolution based on TV regularization (SRTV) [6] extends the classical ROF [8] to seek equilibrium state of TV norm from the magnified image and the fidelity of this image to the noisy input in its original size. Despite the simple implementation in TV, SRTV enhances details of the output images. However, artifacts such as false contours and edge distortions may occur in the resulted images from SRTV.

Glasner’s and Singh’s methods are possible to make a magnified image which is high quality structure rather than interpolation. The magnified image do not occur blur, jaggy and another artifacts. However, these methods do not mention to make fine textures in a magnified image. And high frequency components are not added sufficiently. Therefore, textures of a magnified image are not fine.

In this paper, we propose an approach to make a fine texture magnified image by refining the texture image in SRTV’s generation process. The approach uses Canny edge detection to localize the edge and to remove the non-edge regions in the texture image. Our experimental results show that the proposed approach generates magnified images with fined textures than the SRTV’s and Glasner’s.

2. Super-resolution based on TV Regularization

The SRTV extends ROF to finds $U$, an $M$ magnified $MP \times MQ$ image, where the following equation is satisfied

$$F(U) = TV_U + \lambda \sum_{i,j}^P (s * U^*)_{i,j} - f_{i,j}]^2.$$
Here, $U'_{i,j}$ is the down-sampled $U_{i,j}$ which has the same dimension as $f_{i,j}$, the original low-resolution image. $s$ and $\lambda$ represent the down-sampling and equilibrium parameters. $P$ and $Q$ are the number of rows and columns, respectively. Since the resulted $U$ is a skeleton image, subtracting down-sampled $U$ from $f_{i,j}$ will produce a texture image. Finally, SRTV obtains super-resolution by combining the skeleton and the linearly up-sampled texture images. Figure 1 shows the generation process of super-resolution using SRTV.

In [6], learning-based method was introduced to cope with the issues of artifacts in the resulted images. Although this approach may improve the results, the results vary depending on the chosen images used for learning.

3. Our proposed approach

To enhance SRTV, we focused on how to create high quality texture images. We consider that the high frequency components in an image are taking part in regions of edge

---

Figure 2: Texture images generated from high- and low-resolution images (SIDBA: Pepper)

Figure 3: Steps for non-edge removal process
structures and fine textures. We considered that these regions can be derived using ROF by changing the value of \( \lambda \). Figure 2 shows high-resolution texture images generated from (a) high- and (b) low-resolution images of the standard image database (SIDBA) [9], respectively. \( \lambda \) ranges from -1.0 to 1.0 was used to generate texture images with different details. Figure 2(b) shows that the smaller \( \lambda \) (\( \lambda=0.5 \)) highlights the edge structures while the bigger \( \lambda \) (\( \lambda=0.2 \)) exposes the fine textures. Here, we can observe that the resulted texture images of (a) and (b) are comparable.

Although decreasing the value of \( \lambda \) to some extent may produce a texture image with clear edge structures, it would be difficult to find the optimal value of \( \lambda \) to preserve the edge and in the same time to eliminate the non-edge regions. In this study, we simply remove the non-edge regions by defining the edge regions using an edge detector. The processing steps to remove the non-edge regions are briefly described as follows:

**Step 1** Generate a texture image \( V_{i,j} \) with a small \( \lambda \).
**Step 2** Using SRTV, resize the original image \( f_{i,j} \) to \( g_{i,j} \) with the same dimension as \( V_{i,j} \).
**Step 3** Generate an edge image \( e_{i,j} \) from \( g_{i,j} \) using Canny edge detector.
**Step 4** Referring to every pixel of \( e_{i,p} \) remove pixels at \((i, j)\) in \( V_{i,j} \) when there are no edge pixels found within a \( n \times n \) pixel neighborhood window.
**Step 5** Stop the process if the [Step 4] has been applied to all pixels in \( V_{i,j} \).

Figure 3 illustrates the process of removing non-edge regions from a texture image. All pixels in the texture image were evaluated against those in the corresponding edge image calculated by Canny detector (\( \sigma = 2.0 \), gradient strength \( \geq 3 \), kernel = 10 x 10 pel.). An 8 x 8 window was used in the calculation process. The resulted texture image shows that the process was effective on removing the non-edge regions while preserving the important edge’s. However, the window size for the calculation may need to be adjusted according to the texture complexity of the input image.

To generate a texture image with better edge structures, we introduce a hybrid texture image \( H_{i,j} \) by merging texture images generated with multiple values of \( \lambda \). We calculate \( H_{i,j} \) as follows:

\[
H_{i,j} = \frac{V_{i,j}^{\lambda=0.3} + V_{i,j}^{\lambda=0.2}}{2} \quad \text{(for edge regions)}
\]

\[
H_{i,j} = \frac{V_{i,j}^{\lambda=0.5} + V_{i,j}^{\lambda=0.4} + V_{i,j}^{\lambda=0.3} + V_{i,j}^{\lambda=0.2}}{4} \quad \text{(otherwise)}.
\]

Here, \( V_{i,j}^{\lambda} \) represents an image texture \( V_{i,j} \) generated with \( \lambda \). The determination for edge and non-edge regions follows steps as described above. The value of \( \lambda \) ranges from -0.5 to -0.2. Although this value was determined empirically, for most images, \( \lambda \) ranges from -0.3 to -0.2 will generate texture images with remarkable resemblance to the edge structures of the input image. Conversely, \( \lambda \) ranges from -0.5 to -0.4 will generate texture image with fine textures.

Figure 4 shows the framework of our approach. The difference between SRTV and ours is that the proposed approach generates a hybrid texture from multiple textures calculated with multiple values of \( \lambda \).

**4. Results**

To evaluate the performance of our approach, we generated four and three times magnification of images from those being used in [4]. Figure 5 and 6 show the resulted images generated by [4], SRTV, and our approach, respectively. Both images of SRTV and ours are calculated 200 and 30 iterations to minimize the total variation.

As shown in figure 5 and 6, the differences of perceptual qualities among the magnified images are significant. Our results show finer texture than those processed by [4] and SRTV, especially for the regions of hat (figure 5), koala’s fur and bole (figure 6).

We observed over-sharpening in flat areas of our images appears as noises and false contours at the border between flat and edge areas. These problems arose as we
assigned negative values of $\lambda$ during the calculation of ROF to generate the hybrid texture images. However, due to these values, we are able to obtain texture images to fine texture in a magnified image. We will propose a texture image for flat areas to eliminate flat area’s noise.

5. Conclusion

In this paper, we have presented an approach to generate super-resolution from a single image by composing a hybrid texture image during iterative total variation regularization. The multiple values of $\lambda$ used during the calculation of ROF enable the hybrid texture image to reconstruct the possible high frequency texture for the magnified images. The proposed approach was shown to be able to generate finer images than those of the existing methods. Further improvement may be done in distinguishing between texture and noise, and in avoiding the occurrences of over-sharpening.
References


Figure 6: 3x magnified images processed by each approach