Multiple Inpainting of low resolution images using examplar and super resolution algorithm

A. Breezy Joy, C. Rekha

Abstract—Inpainting is the process of filling the missing regions in an image. The main aim of this paper is to fill the missing areas using examplar based inpainting and to recover the missing areas and improve the quality of the image using super resolution algorithm. The performance of the algorithm is evaluated using PSNR, mean square error and Histogram error. The damaged image is first downsampled. Then the image is inpainted a number of times using examplar based approach. The inpainted images are combined using loopy belief propagation. The image is finally upsampled and the image quality is improved using a super resolution algorithm.

Index Terms—examplar based inpainting, mean square error, loopy belief propagation, running time.

1. INTRODUCTION

Inpainting is used for reconstructing lost or deteriorated parts of images. Image inpainting techniques are in use over a long time for various applications like scratch removal, restoring the damaged, missing portions or removal of objects from the images, etc. In photography and cinema, inpainting is used for film restoration. To reverse the deterioration, dust spots in film and see infrared cleaning. It is also used for removing red-eye, eliminating the stamped date from photographs and removing objects for creative effect. This technique is be used to replace the lost blocks in the transmission and coding of images such as in a streaming video. It is also used to remove the logos in videos. If any unwanted portion of the image is to be removed, then the missing region can be filled using inpainting.

In a paper fragment based image completion, adaptive neighborhood algorithm is used. Here the most similar fragments from the neighborhood are selected and filled in the missing regions. The visible part of the image serve as a training set for the unknown parts. But in this method, more time is spent for searching the best matching fragments. Sometimes two fragments of same part may differ under slight illumination changes. The next paper is inpainting by patch propagation using patch sparsity. In this method the missing patches are represented by sparse linear combination of candidate patches. The patch with large structure sparsity is used for inpainting. Here only the structure of the patch is considered. Texture is not included.

Examplar based inpainting based on local geometry uses K-nearest neighbor algorithm to find the structure tensor. To find the best matching patch, template matching is performed. This method fails to replicate texture in a coherent manner. In texture synthesis method, the textural information and structural information are propagated and the patch priority is calculated. This method fails to handle curved structures. If similar patches do not exist, it does not produce reasonable output.

In a paper statistics of patch offsets for image completion, the statistics of offsets which are sparsely distributed are obtained. Some dominant offsets provide reliable information for filling the image. A stack of shifted images are combined by optimization and it is used to fill missing region. Dominant offsets show how the patterns are repeated and give details about the missing regions. The main advantage is this method has better results and has fast image completion. But this method fails when the desired offset does not have dominant statistics. This problem can be solved only partially by introducing manual offsets.

In non-parametric sampling, a markov random field model is assumed and the conditional distribution of each pixel is calculated. A number of textures are synthesized from a single texture. But some textures may slip into wrong part and produce verbatim copies of the original. Only frontal parallel textures are handled. Local linear embedding is used to generate a high resolution image from a low resolution image. Here more number of neighbours should be considered. All these above methods have low signal to noise ratio and image quality is low. These drawbacks are overcome by the proposed method.

The proposed method uses examplar based inpainting. Inpainting is performed on a coarse version of the input image. For filling the missing regions a non-parametric patch sampling method is used. In this method the best matching patches from the neighborhood are chosen and they are filled in the missing region. To recover fine details on the missing areas, super resolution algorithm is used which converts the low resolution image back to high resolution image. Template matching, non local means and non negative matrix factorization. These three methods aimed at selecting the best matching pixel which is to be replaced in the place of inpainting region. The parameters like PSNR, MSE and Histogram error are estimated.

2. PROPOSED METHOD

An image to be inpainted is taken. As low resolution images are easier and efficient for inpainting, the image is first downsampled. i.e., a high resolution image is converted into low resolution image. Then the image is inpainted for a number of times using examplar based inpainting. Then all the inpainted images are combined using loopy belief propagation. Then to improve the quality of the image, a super resolution algorithm is applied on the image to improve image quality and then convert the low resolution image back to high resolution images. Thus the final inpainted image with high resolution is obtained and the performance is evaluated.
3. EXEMPLAR BASED INPAINTING

The first step in image inpainting is to convert the high resolution image into low resolution image. For this converting process a Gaussian pyramid decomposition is performed. A low pass filtering operation is performed on the image pixels. Downsampling is necessary because the low resolution images are easier to inpaint. Low resolution images are contaminated by noise. Local orientation singularities which affect the filling process are strongly reduced if low resolution images are used.

After converting into low resolution images, mark the target region $\Omega$ to be filled using region of interest. The remaining part of the image is the source region from which the neighbouring patches are selected.

$$\phi = I - \Omega$$

A Boolean matrix is constructed which stores a binary value 1 for pixel to be inpainted and binary value 0 for the other pixels. This matrix is called fill region. The next step is to find the boundary of marked region $\partial \Omega$ by convolving the fill region with laplacian filter. The next step is to find the priority of the patch which is to be filled first. Texture synthesis is performed which copies the samples from the neighbourhood and replaced in the missing region.

3.1 FIND PATCH PRIORIT

The missing patch which has the highest priority is to be found. Priority is the patch which has less unknown regions. The patch priority is given by the product of data term and confidence term.

$$P(p) = C(p) D(p)$$

Where
- $P(p)$ is the patch priority
- $C(P)$ is the confidence term
- $D(p)$ is the data term

Data term can be either a tensor based or sparsity based data terms. Tensor based priority term is given by

$$J = \sum_{i=1}^{m} \nabla l_i \nabla l_i^T$$

Where $J$ is the sum of structure tensors of each scalar image channels. The scalar structure tensor undergoes smoothing operation by

$$J_\sigma = J \times G_\sigma$$

Where $G_\sigma$ is the Gaussian distribution with standard deviation $\sigma$. A similarity weight is measured between the known patches and the unknown patch. For any selected patch, a collection of neighbouring patches with highest similarities are distributed in the same structure or texture. The confidence values of structure for a patch are measured by the sparseness of its non-zero similarities to the neighbouring patches. The patch which is more sparsely distributed with non zero similarities is placed on the fill front due to more structure sparseness. For a patch located at the fill front $\partial \Omega$, a neighborhood window $N(p)$ is set with centre $p$.

3.2 TEXTURE SYNTHESIS

Texture synthesis is the process of algorithmically constructing a large digital image by taking advantage of the structural content. Texture synthesis is used for filling the missing parts. The patch which has the highest priority is filled first. To fill a patch, the most similar patch from the remaining part of the image is taken. A similarity metric is used to find the best matching patch.

Texture synthesis is of two types. Pixel based texture synthesis and patch based texture synthesis. In pixel based method, a texture is synthesized in scan line order by finding and copying pixels with most similar local neighbourhood as the synthetic texture. This method is very useful for image completion.
In patch based texture synthesis, a new texture is created by copying and stitching the textures at various offsets. This method is more effective and faster than pixel based texture synthesis. The chosen patch should have maximum similarity between the known pixel values of the current patch to be filled and the co-located pixel values.

Coherence measure is given by

$$\text{coh}(\Psi_{p_r}) = \min_{p_j \in x} (d_{SSD}(\Psi_{p_r}, \Psi_{p_j}))$$

$d_{SSD}$ - sum of square differences

$\Psi_{px}$ - current patch

$\Psi_{pj}$ - neighbouring patch

Coherence measure is the degree of similarity between the synthesized patch and the original patch. The chosen neighbor should lie within $$(1 + \alpha)d_{\text{min}}$$ where $d_{\text{min}}$ is the minimum distance between the current patch and the closest neighbor.

If the selected patch does not satisfy this condition, then the filling process is stopped and the priority of the current patch is decreased. Again the patch with highest priority is chosen. The unknown parts are pasted using the best matching patch by direct sampling. Alpha blending is used to combine the known part and the source patch. Alpha blending is the process which displays a bitmap which has transparent or semi-transparent pixels. The image is made transparent or the known part and unknown part are combined together. Poisson fusion is the process used to hide the seams. The unevenness caused in the image in the inpainting process is corrected by using Poisson fusion. Patch size and filling order must be considered. Patch size can be $5 \times 5$, $7 \times 7$, $9 \times 9$, $11 \times 11$ etc. a parameter $M$ is considered which is the number of times inpainting is done.

### 3.3 LOOPY BELIEF PROPAGATION

LBP is a message passing algorithm. A node is used to pass a message to the adjacent node only when it has received all the messages, eliminating the message from the destination node to itself. In the same way, all the inpainted images are combined together using loopy belief propagation. A label is assigned to each pixel of the unknown regions $T$ of the image. A major drawback of belief propagation is that it is slow when the number of labels is high. So loopy belief propagation is used to avoid this complexity.

To solve the problems like blur and spatial consistency, there is a need to minimize the objective function. The number of labels assigned must be equal to the number of patches in the source region. In loopy belief propagation, the number of labels is small. Label is the index of the inpainted image from which the patches are extracted. A label is assigned for each pixel so that the total energy $E$ of the markov random field is minimized.

Energy of MRF is given by

$$E(l) = \sum vd(lp) + \sum Vs(ln, lm)$$

$vd(lp)$ is the label cost

$Vs(ln, lm)$ is the discontinuity cost

The cost increases when the similarity between the current patch and colocated patches are less. Discontinuity cost is the difference between labels. $\lambda$ is the weighing factor and it is set to 100. Using loopy belief propagation, minimization of energy $E$ is performed over the target region $T$ and it corresponds to maximum a posteriori (MAP) estimation problem. When $\lambda$ is 0, there is no smoothness term. Some artifacts are visible. A good trade-off is obtained by setting the value of $\lambda$ to 100.

### 4. SUPER RESOLUTION ALGORITHM

After combining all the inpainted images, a super resolution algorithm is used to recover back the high resolution image from the low resolution image. The downsampled image is upsampled. It is also used to improve the quality of the image. In low resolution image, the pixel density within an image is small. Hence it offers fewer details. In a high resolution image, the pixel density is larger. Hence it offers more details. Super resolution is the process of obtaining a high resolution image to low resolution image. To increase the image resolution, the pixel size is reduced by increasing the number of pixels per unit area. But the amount of light available per pixel also decreases. The chip size can be reduced but the increase of capacitance leads to storage problem. SR image reconstruction is computationally effective and cost is less.

### 5. ALGORITHM FOR INPAINTING

- Read an image and find the size of the image. Using median filter, filter the image which replace the pixel value with the neighbourhood of its value.
- Convert the RGB colour image into gray scale image. Create a mask which is completely filled with 0 and the area to be inpainted is marked as 1.
- Create a label matrix which indicates the label for the connected components and used to mark the part of region to be inpainted.
- To set the boundary $(Y_{\text{max}}, Y_{\text{min}}, X_{\text{max}}, X_{\text{min}})$ in the inpaint image, compute the values using the below mentioned equations
  1. $Y_{\text{min}}=\min\text{(row)}-1$
  2. $Y_{\text{max}}=\max\text{(row)}+1$
  3. $X_{\text{min}}=\min\text{(column)}-1$
  4. $X_{\text{max}}=\max\text{(column)}+1$
- Calculate the nearest matching pixel in column and row,

  Matching pixel row, $A= (Y_{\text{max}}-Y_{\text{min}})+1$

  Matching pixel column, $B= (X_{\text{max}}-X_{\text{min}})+1$

- Find the bottom, top, left and right area of the image,  

  1. Bottom Area Range (BAR) & Bottom Mean

     If $(Y_{\text{max}}+A) <= \text{width of the image}$, then $\text{BAR}= (Y_{\text{min}}+A)$ to $(Y_{\text{max}}+A)$ in $x$ direction
     
     $=X_{\text{min}}$ to $X_{\text{max}}$ in $y$ direction.

     Otherwise,

     $\text{BAR} = (Y_{\text{min}}+A)$ to width of image in $x$ direction

     $=X_{\text{min}}$ to $X_{\text{max}}$ in $y$ direction.
\[ \sum_{1,2,\ldots,n} \text{pixels} / \text{(number of pixels)} = \text{mean} \left( \text{BAR} \right) \]

2. Top area range (TAR) & top mean
If \( (Y_{\text{min}} - A) \geq 1 \), then

\[ \text{TAR} = (Y_{\text{min}} - A) \text{ to } (Y_{\text{max}} - A) \text{ in } x \text{ direction} \]

\[ \text{TAR} = X_{\text{min}} \text{ to } X_{\text{max}} \text{ in } y \text{ direction}. \]

Otherwise,

\[ \text{TAR} = 1 \text{ to } (Y_{\text{max}} - A) \text{ in } x \text{ direction} \]

\[ = X_{\text{min}} \text{ to } X_{\text{max}} \text{ in } y \text{ direction}. \]

\[ \sum_{1, 2, \ldots, n} \text{pixels} / \text{(number of pixels)} = \text{mean(TAR)} \]

3. Right area range (RAR) & right mean
If \( (X_{\text{max}} + B) <\!\!\!\!\!\!\!\!\!\!\!\!\!\!(\text{height of the image}) \), then,

\[ \text{RAR} = Y_{\text{min}} \text{ to } Y_{\text{max}} \text{ in } x \text{ direction} \]

\[ = (X_{\text{min}} + B) \text{ to } (X_{\text{max}} + B) \text{ in } y \text{ direction}. \]

Otherwise,

\[ \text{RAR} = Y_{\text{min}} \text{ to } Y_{\text{max}} \text{ in } x \text{ direction} \]

\[ = (X_{\text{min}} - B) \text{ to height of image in } y \text{ direction}. \]

\[ \sum_{1, 2, \ldots, n} \text{pixels} / \text{(number of pixels)} = \text{mean(RAR)} \]

4. Left area range (LAR) & left mean
If \( (X_{\text{min}} - A) \geq 1 \), then,

\[ \text{LAR} = Y_{\text{min}} \text{ to } Y_{\text{max}} \text{ in } x \text{ direction} \]

\[ = (X_{\text{min}} - B) \text{ to } (X_{\text{max}} - B) \text{ in } y \text{ direction}. \]

Otherwise,

\[ \text{LAR} = Y_{\text{min}} \text{ to } Y_{\text{max}} \text{ in } x \text{ direction} \]

\[ = 1 \text{ to } (X_{\text{max}} - B) \text{ in } y \text{ direction}. \]

\[ \sum_{1, 2, \ldots, n} \text{pixels} / \text{(number of pixels)} = \text{mean(LAR)} \]

- In template matching method, the maximum of the mean values of bottom, top, left and right area are calculated. The area to be inpainted is replaced by the maximum of the above four mean values.
- In NLM method, if \( (\text{bottom mean-top mean}) \) is greater than 20, assume the top to bottom value as 0, otherwise 1.
- If \( (\text{left mean-right mean}) \) is greater than 20, then the left to right value is 0, otherwise 1.
- If both top to bottom and left to right values are 1, then the inpainting area is replaced by \( \text{pixel} = \text{mean(bottom, top, right, left)} \);
- If top to bottom value is 1 and left to right value is 0 then the inpainting area is replaced by \( \text{pixel} = \text{mean(bottom mean, top mean)} \)
- If the top to bottom value is 0 and left to right value is 1, then the inpainting region is replaced by \( \text{pixel} = \text{mean(right mean, left mean)} \)
- If both top to bottom value and left to right are 0, and \( (\text{left mean-right mean}) > (\text{bottom mean-top mean}) \), then the inpainting region is replaced by \( \text{pixel} = \text{mean(bottom mean, top mean)} \)
- Otherwise the pixels are replaced by \( \text{pixel} = \text{mean(right mean, left mean)} \)
- In NMF method, consider a temporary value. If the miss image is less than 10, the inpainting region is replaced by \( \text{Missing pixel} = \text{max temporar} \text{y value} \)
- Finally the image has to be enhanced. Median filter is used to filter the images obtained from the three methods and hence the image quality is increased.
- The performance of inpainting is evaluated by calculating PSNR mean square error and histogram error.

6. RESULTS AND DISCUSSION

The final output consists of the input image, inpainted image, TM image, NLM image and NMF image. First the input image is displayed. Next the binary mask is displayed in which the region to be inpainted is in white colour and the remaining portion is displayed as black. The region to be inpainted has the pixel value 1 and the remaining part has the pixel value 0. Three methods of inpainting are performed and their outputs are displayed. The image quality is improved and the downsampled image is converted back to high resolution image. The three methods include Template matching (TM), non local means (NLM) method and non negative matrix factorization (NMF) method. Output of each method is displayed.

Then signal to noise ratio is measured for the three images. But it does not have much variation. Next mean square error is calculated. MSE also does not have considerable changes among the three images. Finally histogram error is calculated. Histogram error is less in TM image than the other two images. Hence experimental results show that among the three methods, TM image is better. Three parameters, signal to noise ratio, mean square error and Histogram error are estimated to evaluate the performance of the three methods. Template matching is better than the other two methods.

Four examples are considered. For baby image, leaf image, nature image and blue image TM, NLM and NMF methods of inpainting are performed. The table 6.1, shows that the PSNR and MSE values are nearly equal and does not show much changes for the three methods. To find which method is better we have to go for another parameter. Histogram error is calculated for TM, NLM and NMF methods. Histogram error is less for TM image as compared to the other two methods. Template matching is the process of selecting the best matching pixel by taking the mean of the top, bottom, left and right value. As the inpainting area is replaced by the mean of the pixel, the inpainting process will be better. Hence TM method is suitable for inpainting. The comparison of PSNR, MSE and Histogram error is shown in the table below. The output figures of each method with different examples are shown below.
6.1 Comparison of PSNR, MSE and Histogram Error

<table>
<thead>
<tr>
<th>Image name</th>
<th>Method</th>
<th>PSNR</th>
<th>MSE</th>
<th>Histogram Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baby</td>
<td>TM</td>
<td>49.3500</td>
<td>0.7552</td>
<td>2.7458e-005</td>
</tr>
<tr>
<td></td>
<td>NLM</td>
<td>49.2528</td>
<td>0.7723</td>
<td>3.1116e-005</td>
</tr>
<tr>
<td></td>
<td>NMF</td>
<td>49.3539</td>
<td>0.7545</td>
<td>2.8679e-005</td>
</tr>
<tr>
<td>Leaf</td>
<td>TM</td>
<td>30.4311</td>
<td>26.8806</td>
<td>2.3023e-004</td>
</tr>
<tr>
<td></td>
<td>NLM</td>
<td>30.4109</td>
<td>26.1553</td>
<td>3.1217e-004</td>
</tr>
<tr>
<td></td>
<td>NMF</td>
<td>30.4445</td>
<td>26.6987</td>
<td>3.2135e-004</td>
</tr>
<tr>
<td>nature</td>
<td>TM</td>
<td>38.0800</td>
<td>10.1177</td>
<td>8.7285e-005</td>
</tr>
<tr>
<td></td>
<td>NLM</td>
<td>38.0730</td>
<td>10.1340</td>
<td>8.9937e-005</td>
</tr>
<tr>
<td></td>
<td>NMF</td>
<td>38.0841</td>
<td>10.1082</td>
<td>9.2380e-005</td>
</tr>
<tr>
<td>blue</td>
<td>TM</td>
<td>36.8829</td>
<td>13.3287</td>
<td>2.1635e-004</td>
</tr>
<tr>
<td></td>
<td>NLM</td>
<td>36.8840</td>
<td>13.3254</td>
<td>2.9234e-004</td>
</tr>
<tr>
<td></td>
<td>NMF</td>
<td>36.8841</td>
<td>13.3252</td>
<td>2.4539e-004</td>
</tr>
</tbody>
</table>

Table 6.1 Comparison of PSNR, MSE and Histogram Error

![Original image](image1.png) ![damaged image](image2.png) ![inpainting region](image3.png)

![TM image](image4.png) ![NLM image](image5.png) ![NMF image](image6.png)
7. CONCLUSION

Thus the inpainting was done for a series of images using examplar based inpainting. First the input image is downsampled to make the inpainting easier. Then inpainting is done for a number of times using three methods of inpainting. They are template matching (TM), non local means (NLM) and non negative matrix factorization (NMF). Finally a super resolution algorithm is used to improve the quality of the image. It undergoes a filtering operation to reduce the noise in the image. Among the three methods, the PSNR and mean square error does not show much variation. But histogram error is low for TM image as compared to the other two methods. Hence template matching is the best method suitable for inpainting. Future work would be to extend this to other super resolution methods and also to improve the image quality after inpainting.

REFERENCES


