New Index Generation Method for Example-Based Super-Resolution using DCT Signs

Takaharu Kouda*a

* Department of Electrical Engineering, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, JAPAN

Abstract

The purpose of this study is to propose a new method for generating indexes for example-based Super-Resolution. Conventional method had a problem that it takes a long time to search for high-frequency component from the database. The proposed method aims to accelerate the database search by generating an index expressing patterns of gray level in the local block from DCT signs. First, the low-resolution input image is expanded using bicubic interpolation. Local blocks are extracted from the enlarged image and the index is generated. The optimal high frequency components for the local block are then searched for from the database using the generated index. The enlarged image is then converted into a high definition image by adding the searched high frequency components to the local blocks. As a result, a high-definition enlarged image can be obtained at faster speeds using the proposed method than with the conventional method. In conclusion, we have demonstrated that the DCT sign can adequately express gray level patterns in local blocks and is effective in speeding up example-based Super-Resolution.

Keywords: Example-based super-resolution, DCT sign, index.

1. Introduction

In recent years, the use of large, flat-screen televisions has spread to ordinary households, enabling us to enjoy video on high-resolution display devices. However, the video shot for DVDs, etc., is of a lower resolution. Thus, it is necessary to find a method of converting such video to a higher resolution when it is displayed on large-screen televisions.

A simple way of converting low-resolution images to high-resolution images is the use of image interpolation methods such as linear interpolation and bicubic interpolation. However, such image interpolation methods perform interpolation with approximate values between pixels when converting the input image into a high-resolution image. Due to this, such methods may be unable to restore the high frequency components lost at the point of observation, leading to a blurred image. For this reason, super-resolution techniques that improve the resolution of the captured image have been gaining attention.

Super-resolution is a technique that improves the sense of an image’s resolution by generating high frequency components that do not exist in the original image when converting the input image into a high-resolution image. Techniques for super-resolution include both single-frame super-resolution and multi-frame super-resolution. This study proposes a new method for Example-Based Super-Resolution within single frame super-resolution.

A typical method of example-based Super-Resolution is Image Hallucination[1,2]. This method first expands the low-resolution image to the objective resolution using bicubic interpolation. This enlarged image is referred to as a low frequency image. Next, a 9x9 [pixel] local block is cut from the contour portion of the low frequency image. A pre-learned database is searched for high frequency components that are then added to the extracted local block. A high-resolution image is obtained by applying the above process to the entire low frequency image.

However, the conventional method presented an issue in regards to slow processing speeds, which was a result of the large amount of time taken to search for the optimal high frequency components for the local block from the database.

The purpose of this study is to propose a new method
for generating indexes for example-based Super-Resolution. The proposed method aims to accelerate the database search by generating an index expressing patterns of gray level in the local block from DCT signs.

First, the low-resolution input image is expanded using bicubic interpolation. Local blocks are extracted from the enlarged image and the index is generated. The optimal high frequency components for the local block are then searched for from the database using the generated index. The enlarged image is then converted into a high definition image by adding the searched high frequency components to the local blocks.

As a result, a high-definition enlarged image can be obtained at faster speeds using the proposed method than with the conventional method.

In conclusion, we have demonstrated that the DCT sign can adequately express gray level patterns in local blocks and is effective in speeding up Example-Based Super-Resolution.

2. Proposed Method

In the proposed method, a high-definition enlarged image is obtained by adding the high frequency component to the image that is enlarged by bicubic interpolation. First, the low resolution input image is expanded using bicubic interpolation. These are referred to as low frequency images in this study. Next, the database is searched for the optimal high-frequency components for the 8×8[pixel] local blocks in the low frequency image. By applying this process to the entire low frequency image, the high-frequency components lacking in the low frequency image are generated. High definition enlarged images are obtained by adding these high-frequency components to the low frequency image.

2.1 DCT sign index

The DCT are commonly used in the field of image compression and pattern recognition. DCT and inverse DCT are defined as follows:

DCT

\[
f(m, n) = \sum_{u=0}^{W-1} \sum_{v=0}^{H-1} C(u) C(v) F(u, v)
\times \cos \left(\frac{(2m+1)u\pi}{2W}\right) \cos \left(\frac{(2n+1)v\pi}{2H}\right)
\]  

where,

\[
C(w) = \begin{cases} 
\frac{1}{\sqrt{2}} & (w = 0) \\
1 & (w \neq 0)
\end{cases}
\]

Here, \(w\) is \(u\) or \(v\). Inverse transformation in the case of DCT signs only is referred to as Sign-only Synthesis (SOS). The DCT SOS is defined as

\[
f(m, n) = \sum_{u=0}^{W-1} \sum_{v=0}^{H-1} C(u) C(v) e^{i\theta}
\times \cos \left(\frac{(2m+1)u\pi}{2W}\right) \cos \left(\frac{(2n+1)v\pi}{2H}\right)
\]

where,

\[
\theta = \begin{cases} 
0 & F(u, v) \geq 0 \\
\pi & F(u, v) < 0
\end{cases}
\]

Figure 1 shows the sample of SOS processing. Looking at figure 1, we can see that SOS detects the edge of the

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>11</td>
<td>10</td>
<td>16</td>
<td>24</td>
<td>40</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>14</td>
<td>19</td>
<td>26</td>
<td>58</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>16</td>
<td>24</td>
<td>40</td>
<td>57</td>
<td>69</td>
</tr>
<tr>
<td>14</td>
<td>17</td>
<td>22</td>
<td>29</td>
<td>51</td>
<td>87</td>
<td>80</td>
</tr>
<tr>
<td>18</td>
<td>22</td>
<td>37</td>
<td>56</td>
<td>68</td>
<td>109</td>
<td>103</td>
</tr>
<tr>
<td>24</td>
<td>35</td>
<td>55</td>
<td>64</td>
<td>81</td>
<td>104</td>
<td>113</td>
</tr>
<tr>
<td>49</td>
<td>64</td>
<td>78</td>
<td>89</td>
<td>103</td>
<td>121</td>
<td>120</td>
</tr>
<tr>
<td>72</td>
<td>92</td>
<td>95</td>
<td>98</td>
<td>112</td>
<td>100</td>
<td>103</td>
</tr>
</tbody>
</table>

Fig. 1. Sample of Sign-only Synthesis.

Fig. 2. JPEG quantization table.
image. From this, we can say that the DCT coefficient signs contain image edge information. Therefore, the DCT sign is used such as pattern recognition[4,5]. In this study, to express gray patterns in the local blocks of the image, a unique index is generated using the DCT coefficient signs of local blocks.

The processing order is as shown below. First, DCT is carried out on the 8×8[pixel] local blocks of the image. Fifteen DCT coefficient signs, excluding DC component, are extracted from the DCT coefficient low-frequency area 4×4[pixel]. The DCT signs extracted at this time are 1 if the DCT coefficient is 0 or more. Otherwise, they are 0. A 15-bit integer is then generated from the extracted DCT signs. Furthermore, using the JPEG quantization table shown in Figure 2, the amplitude of the low-frequency area 2×2[pixel] of DCT coefficient are thresholded. Where, DC component is excluded from its threshold processing. The amplitude information of the low frequency area is extracted as 1 when the DCT coefficient amplitude exceeds the threshold value and 0 when it does not. An 18-bit integer is generated by adding 3-bit amplitude information to the above 15-bit integer. This 18-bit integer is referred to as the DCT sign index.

2.2 Database creation

Here, a database is created using multiple natural images as learning image. First, after convoluting a distribution 1.0 Gaussian filter for the learning image, the low-resolution image is created using down-sampling. Next, the low-resolution image is expanded to the original resolution using bicubic interpolation. This is a low-frequency image corresponding to the learning image. The DCT sign index is generated after extracting the 8×8[pixel] local blocks from the low frequency image. Next, the differences between the learning image and low frequency image are calculated for each local block. These are the high-frequency components for the local blocks. These high-frequency components are registered in the database using the DCT sign index.

2.3 Super-Resolution processing

First, the input image is expanded using bicubic interpolation. The 8×8[pixel] local blocks are extracted from the enlarged image and the DCT sign index is generated. Next, the optimal high frequency component for the local block is searched for from the database using the DCT sign index. The extracted high-frequency components are output in the high frequency image as a weight in the

![Fig. 3. Simulation result 1.](attachment:image.png)
Gaussian function. By applying this process to the whole of the image, the high-frequency components lacking in the enlarged image are generated. The final generated high frequency image is synthesized with the enlarged image. However, the high-frequency components may occur as noise in the synthesized image. In this case, noise that was not in the input image is removed using the back projection method. The final processing results are obtained after performing the above process.

### 3. Simulation

Figure 3 and Figure 4 shows the processing results using the conventional method and the proposed method. These figures show the result of expanding the original image two times. The original image size is $256 \times 256$ [pixel]. PSNR is calculated using a correct image of $512 \times 512$[pixel]. Table 1 shows the processing speeds of the conventional method (Image Hallucination) and the proposed method, and Table 2 shows the experimental environment used at the time. Table 1 shows that with the proposed method, processing complete within 1.0 second if the image size is under 350,000 pixels.

From the simulation results, it was found that the result of the proposed method is smooth contours than the conventional method. We can see that the PSNR of the proposed method also improved when compared with the results of the conventional method. Furthermore, when the input image had 1.2 million pixels, the processing speed was approximately 1/60 that of the conventional method.

![Figure 3: Bicubic (PSNR = 27.506 [dB]).](image)

![Figure 4: Image Hallucination (PSNR = 30.434 [dB]).](image)

![Figure 5: Proposed method (PSNR = 31.544 [dB]).](image)

**Table 1.** Processing time [sec.].

<table>
<thead>
<tr>
<th>Resolution [pixel]</th>
<th>Proposed method</th>
<th>Image Hallucination</th>
</tr>
</thead>
<tbody>
<tr>
<td>$256 \times 256$</td>
<td>0.124</td>
<td>30.487</td>
</tr>
<tr>
<td>$512 \times 512$</td>
<td>0.498</td>
<td>51.694</td>
</tr>
<tr>
<td>$720 \times 480$</td>
<td>0.638</td>
<td>66.843</td>
</tr>
<tr>
<td>$960 \times 720$</td>
<td>1.296</td>
<td>106.500</td>
</tr>
<tr>
<td>$1280 \times 960$</td>
<td>2.304</td>
<td>132.953</td>
</tr>
</tbody>
</table>

**Table 2.** Experimental environment.

<table>
<thead>
<tr>
<th></th>
<th>Windows XP Professional SP3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td>Intel Xeon W3570 3.2GHz</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>4GB</td>
</tr>
</tbody>
</table>
4. Conclusions

The purpose of this study is to propose a new method for generating indexes for example-based Super-Resolution. The proposed method aims to accelerate the database search by generating an index expressing patterns of gray level in the local block from DCT signs. First, the low-resolution input image is expanded using bicubic interpolation. Local blocks are extracted from the enlarged image and the index is generated. The optimal high frequency components for the local blocks are then searched for from the database using the DCT sign index. The enlarged image is then converted into a high definition image by adding the searched high frequency components to the local blocks.

As a result, a high-definition enlarged image can be obtained at faster speeds using the proposed method than with the conventional method.

In conclusion, we have obtained the following from this research:

1. Using the DCT sign index, it is possible to search the database for the optimal high frequency components for the local block.
2. High quality enlarged images can be obtained than with the conventional method. Furthermore, they can be processed at faster speeds than with the conventional method.

Future issues include how to more reduce of the processing time in order to process moving images in real-time.

References