Real-Time Interpolation Using Bilateral Filter for Image Zoom or Video Up-Scaling/Transcoding

Kwok-Wai Hung, Student Member, IEEE and Wan-Chi Siu, Senior Member, IEEE
Dept. of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong

Abstract-- In this paper, we propose a novel image interpolation technique using the small-kernel bilateral filter. The range distance of the bilateral filter is estimated using a novel maximum a posterior (MAP) estimation, in order to consider both the diagonal and vertical-horizontal correlations in image. The proposed method requires a few arithmetic operations, such that it is suitable for real time image zoom, video up-scaling and video transcoding applications in HDTV systems. Subjective evaluation shows that the proposed approach outperforms the conventional method in terms of edge sharpness and edge smoothness.

I. INTRODUCTION
The resolution of image or video is often required to be increased in HDTV applications, such as zooming for a better perception, up-scaling to display on high-resolution devices and transcoding low-resolution contents to broadcast on high-definition channels, etc. Image interpolation can reproduce a high-resolution (HR) image from its low-resolution (LR) counterpart. Conventional polynomial-based interpolation methods such as bilinear and bicubic interpolation suffer from annoying artifacts such as aliasing effect, blur, halo, etc, around the edges due to their non-adaptive properties. To address the problem, some edge-directed interpolation methods have been developed. The idea is based on the edge property – the geometric regularity [2]. Yang et al. and Wang et al. explicitly estimated the edge orientation and then interpolated along the edge orientation [3-4]. However, the estimation accuracy of the edge orientation is concerned. Li and Orchard made use of the geometric duality property to propose a least squares estimation method [5]. Zhang and Wu extended Li’s method using soft-decision estimation (called SAI) to consider consistency within a local region [6]. Zhang proposed to fuse two directionally interpolated results using the linear minimum mean square error estimation (LMMSE) [7]. Hung and Siu proposed a fast video interpolation method using the linear motion model [8].

In this paper, we propose a fast edge-directed interpolation algorithm using the small-kernel bilateral filter, which has the advantages of fast computation and stability. A new maximum a posteriori (MAP) estimation of the range distance (for the bilateral filter) is proposed to incorporate both diagonal and vertical-horizontal correlations by using the adaptive regularization. A post-processing step based on the idea of soft-decision estimation is proposed to constrain the result to consider local statistical consistency. Computational analysis shows that the proposed method requires less than 150 arithmetic operations per pixel and it provides the highest PSNR value among available edge-directed algorithms [5-7].

This work is supported by the Center for Signal Processing, the Hong Kong Polytechnic University and the Research Grant Council of the Hong Kong SAR Government (PolyU 5278/08E).

II. PROPOSED METHOD

A. Bilateral filter for image interpolation

The bilateral filter was initially used for image denoising [9]. Consider that the pixel to be interpolated is equivalent to the pixel to be denoised, we formulate the cost function [10]

$$\min_{H_k} \sum_{k=0}^{K} w_k (H - H_k)^2$$

where $H$ is the HR pixel to be estimated and $\{H_k\}$ are the observed neighboring pixels of $H$. Hence the terms $\{H-H_k\}$ are approximations of the first–order derivative in different scales, which are weighted by $\{w_k\}$ according to their importance. We simplify the weight used by bilateral filter [9],

$$w_k = \exp\left[ -1\frac{H - H_k}{2\sigma} \right]$$

where $\exp(.)$ is an exponential operation and $\sigma$ is the variance. Recent papers [4-7] show that four nearest neighbors of $H$ is sufficient to provide high interpolation quality. Hence, $K$ is set to 4 (2x2 local window). Solving (1), we have

$$\hat{H} = \frac{1}{S} \sum_{k=0,1,2,3} w_k H_k$$

where the normalization factor $S$ is the sum of all $w_k$.

B. Proposed MAP estimation of range distance

Let us denote the range distance in (2) as $R_k = |H-H_k|$, the observed neighboring range distance as $\{R_{k,l}\}$ and the range distance formed by interpolating $H$ as $R_{k,bic}$. The proposed maximum a posterior (MAP) estimation of $R_k$ is as follows

$$\min_{R_k} \sum_{l=0,1,\ldots,7} (R_k - R_{k,l})^2 + \lambda (R_k - R_{k,bic})^2$$

where the regularization factor $\lambda$ is adaptive to the ratio of local vertical-horizontal and diagonal gradients because the bicubic interpolation provides an accurate range estimation in case of vertical or horizontal edge. Solving (4), we have

$$\hat{R}_k = \frac{1}{8} + \lambda \left( \sum_{l=0,1,\ldots,7} R_{k,l} \right)$$

where the estimated range distance $\hat{R}_k$ is then substituted into (2) for calculating the weight for the interpolation equation (3).

C. Post-processing to consider local consistency

Pixel-based optimization (3) ignores the correlation of image structure within a local region. We propose the patch-based MAP estimation to constrain statistical consistency

$$\min_{H_k} \left[ \left( H - \sum_{k=0}^{K} w_k H_k \right)^2 + \sum_{k=0}^{K} U_k \left( H_k - \sum_{i=0}^{3} w_i H_{ki} \right)^2 \right]$$

where $\{H_{ki}\}$ are the neighbors of $H_k$. Figure 1 shows one example illustration of $H$, $\{H_k\}$ and $\{H_{ki}\}$. Hence, the regularization term constrains the whole local region to fit the
interpolation equation (3), i.e. statistical consistency. The regularization factor \( U_k \) is defined as: 
\[
U_k = d(c + w_k)
\] 
where the weight \( w_k \) guides the direction of consistency, the small constant \( c \) is for stabilization, and factor \( d \) depends on the edge gradient since a larger gradient exhibits a stronger consistency. Solving (6), we have
\[
\hat{H} = \frac{\sum_{k=0}^{3} w_k H_k + \sum_{k=0}^{3} U_k w_k H_k - \sum_{i=0}^{3} w_i H_{ki}}{1 + U_0 w_0^2 + U_1 w_1^2 + U_2 w_2^2 + U_3 w_3^2}
\] 
which reduces to (3) when \( U_k = 0 \) if the edge gradient is smaller than a threshold, i.e. non-edge pixels.

D. Computational analysis

Table 1 shows that a small number of arithmetic operations per pixel is required for our interpolation equations (3) and (8).

<table>
<thead>
<tr>
<th>Non edge pixels using (3)</th>
<th>Add</th>
<th>Sub</th>
<th>Mul.</th>
<th>Div.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
<td>33</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Edge pixels using (8)</td>
<td>67</td>
<td>37</td>
<td>45</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 2 PSNR (dB) of the proposed approach and other methods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kodim01-24</td>
<td>29.504</td>
<td>29.514</td>
<td>29.408</td>
<td>29.921</td>
<td>29.925</td>
</tr>
<tr>
<td>Lena</td>
<td>33.939</td>
<td>33.927</td>
<td>33.777</td>
<td>34.707</td>
<td>34.637</td>
</tr>
<tr>
<td>Pepper</td>
<td>32.848</td>
<td>33.202</td>
<td>33.014</td>
<td>33.533</td>
<td>33.735</td>
</tr>
<tr>
<td>Splash</td>
<td>40.413</td>
<td>40.036</td>
<td>40.301</td>
<td>41.061</td>
<td>41.588</td>
</tr>
<tr>
<td>Average</td>
<td>29.639</td>
<td>29.688</td>
<td>29.569</td>
<td>30.065</td>
<td>30.101</td>
</tr>
</tbody>
</table>

III. EXPERIMENTAL RESULTS

All tests were run on an Intel i7 950 system. We compare other methods in the literature, including bicubic interpolation [1], LMMSE [7], NEDI [5] and SAI [6]. We implemented bicubic interpolation and our method using C++ codes, which requires 0.0036 and 0.049 second for interpolating one image. The PSNR values of 28 natural images were evaluated in Table 2. The proposed approach provides the highest average PSNR value. Although the PSNR value of SAI is closed to the proposed approach, SAI requires approximately 40 times more computation due to its complicated formulation [6].

Figure 2 shows portions of interpolated parts in the Kodim05, 13, 20 and 19 images. The proposed approach provides obviously results better than Bicubic in all images in terms of edge smoothness and sharpness. Moreover, it provides a much more stable result than that SAI provides in Kodim19 image.

IV. CONCLUSION

In this paper, we have presented a fast image interpolation method using the bilateral filter. We propose a new maximum a posterior (MAP) estimation of the range distance for the bilateral filter. A post processing step is proposed to constrain statistical consistency within a local region. The post-processing step uses MAP estimation with adaptive regularization to exploit the appropriate value of consistency strength, which is dependent on the edge gradient and the bilateral filter weights. Computational analysis shows that the proposed approach is suitable for real time applications while its interpolation quality outperforms other methods in the literature in terms of PSNR values. Subjective evaluations show the proposed approach gives good results in terms of interpolation stability, edge smoothness and edge sharpness, and is suitable for image or HDTV applications.

REFERENCE


Figure 2 Portions of interpolated Kodim05, 13, 20 and 19 images.