A MODIFIED EDGE DIRECTED INTERPOLATION FOR IMAGES

Wing-Shan Tam[1,2], Chi-Wah Kok[1], Wan-Chi Siu[1]

Department of Electronic and Information Engineering, Hong Kong Polytechnic University[1]
Department of Electronic Engineering, City University of Hong Kong[2]

ABSTRACT

A modification of the new edge-directed interpolation method is presented. The modification eliminates the prediction error accumulation problem with adopting a modified training window structure, and further extends the covariance matching into multiple directions for suppressing the covariance mis-match problem. Simulation results show that the proposed method achieves remarkable subjective performance in preserving the edge smoothness and sharpness among other methods in literature. It also demonstrates consistent objective performance among a variety of images.

1. INTRODUCTION

Image interpolation is a process that estimates a set of unknown pixels from a set of known pixels in an image. High quality interpolated images are obtained when the pixel values are interpolated according to the edges of the original images. A number of edge-directed interpolation (EDI) methods that make use of the local statistical and geometrical properties to interpolate the unknown pixel values are shown to be able to obtain high visual quality interpolated images without the use of edge map [1–6]. The *New Edge-Directed Interpolation* (NEDI) method in [1] models the natural image as a second-order locally stationary Gaussian process and estimates the unknown pixels using simple linear prediction. A covariance of the image pixels in a local block (training window) is required for the computation of the prediction coefficients. Compared to the conventional methods, e.g. the bilinear method or the bicubic method, the NEDI method preserves the sharpness and continuity of the interpolated edges. However, this method considers only the four nearest neighboring pixels along the diagonal edges and not all the unknown pixels are estimated from the original image, which degrades the quality of interpolated image. Moreover, the NEDI method has difficulty in texture interpolation because of the large kernel size, which reduces the fidelity of the interpolated image, thus lower the *peak signal-to-noise ratio* (PSNR) level. Markov random field (MRF) model-based method [2] models the image with MRF model and extends the edge estimation in other possible directions by increasing the number of neighboring pixels in the kernel. MRF model-based method is able to preserve the visual quality of the interpolated edges and also maintain the fidelity of the interpolated image, thus enhances the PSNR level. The more accurate the MRF model, the better the efficiency of the MRF model-based method. The *Improved New Edge-directed Interpolation* (iNEDI) method in [4] modifies the NEDI method by varying the size of the training window according to the edge size and achieves better PSNR performance. The *Iterative Curvature Based Interpolation* (ICBI) method in [6] considers the effects of the curvature continuity, curvature enhancement and isophote contour. By proper weighting between these three effects, the ICBI method produces perceptually pleasant image. However, similar to the iNEDI method, the performance depends on the chosen parameters. This paper will present an improvement of the NEDI method, namely the *Modified Edge-Directed Interpolation* (MEDI), which is an extension of our work in [7]. We proposed a different training window to mitigate the interpolation error propagation problem [7]. Later on, we found the similar training window has been proposed in the *Improved Edge-Directed Interpolation* (IEDI) [3]. The enlarged training window will inevitably increase the interpolation error due to the worsen covariance mis-match problem, thus the interpolation results obtained by IEDI is shown to be worse than that of NEDI in most cases. As a result, we propose to apply multiple training windows to mitigate the covariance mis-match problem. The performance of the proposed method is verified by extensive simulations and comparisons with other EDI based interpolation methods, including the NEDI, the IEDI, the iNEDI and the ICBI methods. The simulation results show that the proposed method generates high visual quality images and demonstrates a highly consistent objective performance over a wide variety of images.

2. REVIEW OF THE EXISTING METHODS

Consider the interpolation of a low-resolution image X (with size $H \times W$) to a high-resolution image Y (with size $2H \times 2W$) such that $Y(2i,2j) = X(i,j)$. This is graphically shown in Figure 1, where the white dots denote the pixels from X. The NEDI method is a two-step interpolation process which first estimates the unknown pixels $Y_{2i+1,2j+1}$ (gray dot in Figure 1(a)), then the pixels $Y_{2i,2j+1}$ and $Y_{2i+1,2j}$ (black dot in Figure 1(b)). The NEDI method makes use of a fourth-order linear prediction to interpolate unknown pixel from the four neighboring pixels, e.g. $Y_{2i+1,2j+1}$ is estimated from $\{Y_{2i,2j}, Y_{2i+2,2j}, Y_{2i+2,2j+2}, Y_{2i,2j+2}\}$ as

$$Y_{2i+1,2j+1} = \sum_{k=0}^{1}\sum_{\ell=0}^{1} \alpha_{2k+\ell} Y_{2(i+k),2(j+\ell)} \qquad (1)$$

To simplify the notations, and without ambiguity, the sixteen covariance values and four cross-covariance values obtained by the four pixels in eq.(1) are enumerated to be $R_{k\ell}$ and r_k with $0 \le k,\ell \le 3$, respectively, as shown by the labels next to the arrows in Figure 1(a). For examples, $R_{03} = E\left[Y_{2i,2j}Y_{2i,2j+2}\right]$ and $r_0 = E\left[Y_{2i,2j}Y_{2i+1,2j+1}\right]$. The optimal prediction coefficients set α can be obtain as [1]

$$\alpha = \mathbf{R}_{yy}^{-1}\mathbf{r}_y, \qquad (2)$$

where $\alpha = [\alpha_0, \cdots, \alpha_3]$, $\mathbf{R}_{yy} = [R_{k\ell}]$ and $\mathbf{r}_y = [r_0, \cdots, r_3]$. The interpolation is therefore locally adapted to $\mathbf{R}_{yy}$ and $\mathbf{r}_y$.

This work was supported by the Research Grants Council of Hong Kong SAR Government under the CERG grant no. PolyU5278/8E(BQ14-F).

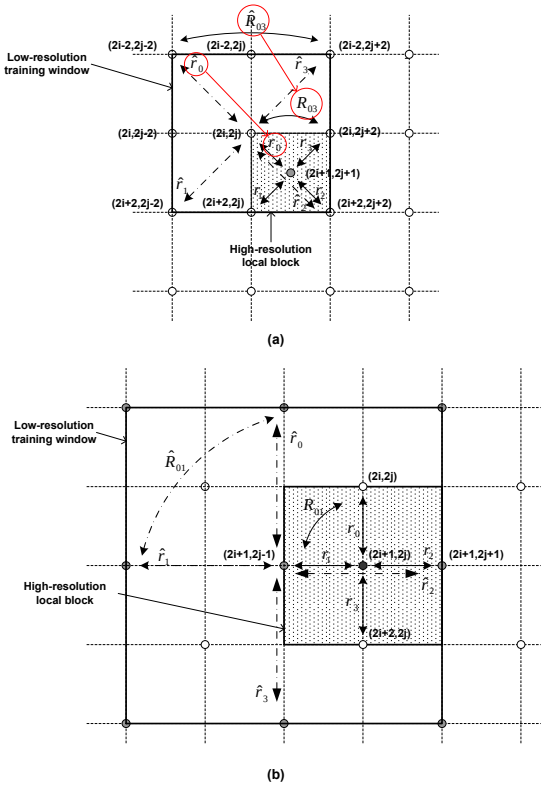

Figure 1: Illustration of the training windows and local blocks of (a) the first step and (b) the second step of the NEDI method.

However, the computation of $R_{k\ell}$ and r_k would require the knowledge of $Y_{2i+1,2j+1}$ which is not available before the interpolation. This difficulty is overcome by the "geometric duality" property, where the covariance $\hat{r}_0$ (circled in the figure) estimated from the low-resolution training window is applied to replace the high-resolution covariance r_0 as indicated by the arrow in Figure 1(a). In a similar manner, the covariance r_k are replaced by $\hat{r}_k$ with $0 \le k \le 3$. The unknown pixel $Y_{2i+1,2j+1}$ is therefore estimated by eq.(1) with $\hat{R}_{k\ell}$ and $\hat{r}_k$. The remaining pixels $Y_{2i,2j+1}$ and $Y_{2i+1,2j}$ can be obtained by the same method with a scaling of $2^{1/2}$ and a rotation factor of $\pi/4$ as shown in Figure 1(b). A hybrid approach is adopted, where covariance based interpolation is applied to edge pixels (pixels near an edge) when the covariance matrix has full rank, and the energy of the covariance matrix is higher than a predefined threshold ε, otherwise bilinear interpolation is applied to non-edge pixels (pixels in smooth regions). However, prediction error is unavoidable in the interpolated pixels. The NEDI method propagates the errors from the first step to the second step because the estimation in the second step depends on the result of the first step (the black dot is estimated from the gray dots as shown in Figure 1(b)). To cater this problem, a modified training window structure has been developed independently in [3, 7]. The training window in the second step of the NEDI method for the interpolation of $Y_{2i+1,2j}$ and $Y_{2i,2j+1}$ is modified to form a sixth-order linear prediction with a 5×9 training window as illustrated in Figure 2, where

$$Y_{2i+1,2j} = \sum_{k=0}^{1} \sum_{\ell=-1}^{1} \alpha_{2k+\ell} Y_{2(i+k),2(j+\ell)}. \qquad (3)$$

The coefficients $\alpha_{2k+\ell}$ can be estimated from eq.(2) with the auto-covariance matrix $\mathbf{R}_{yy}$ that contains thirty-six $R_{k\ell}$, and cross-covariance vector $\mathbf{r}_y$ with six elements of r_k with $0 \le k, \ell \le 5$. The high-resolution covariances are then replaced by the low-resolution covariances of $\hat{\mathbf{R}}_{yy}$ and $\hat{\mathbf{r}}_y$ using geometric duality property. The rest of the unknown pixels $Y_{2i,2j+1}$ can be estimated in a similar manner with a sixth-order linear prediction as that for pixels $Y_{2i+1,2j}$ but with the training window rotated by $\pi/2$.

However, both the NEDI method and the IEDI method suffer from the covariance structure mis-match problem as illustrated in Figure 4, where the white box is the geometric low-resolution training window, the gray box is the corresponding high-resolution local block and the dash lines "AB" and "CD" indicate the image edges in the local block. Figure 4(a) and (b) show the training windows adopted in the NEDI method and the IEDI method. Clearly, the geometric duality property is satisfied as shown in Figure 4(a). However, it is apparent that the geometric duality property is not satisfied for the edge "CD" as shown in Figure 4(b), and thus causes covariance mis-match. To cater this problem, we propose to consider all the four locations of the low-resolution training window and the high-resolution local block as shown in Figure 4(b)–(e).

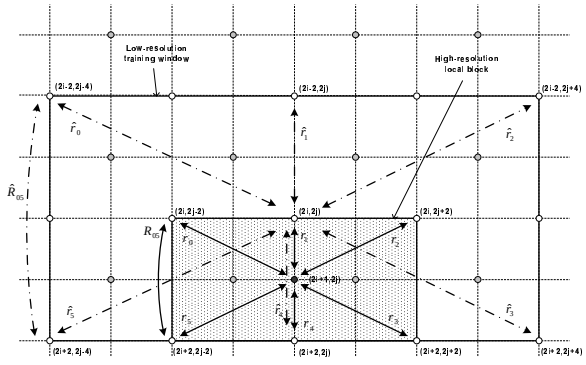

Figure 2: Illustration of the training window and local block of the second step of the MEDI method.

3. THE PROPOSED METHOD: MEDI

To reduce the covariance mis-match problem, multiple low-resolution training windows are used. Figure 4(b–e) illustrates the four training windows applied in the first step of the proposed method. The NEDI and the IEDI methods consider the training window shown in Figure 4(b) only and the training window is centered at pixel $Y_{2i,2j}$. Compared with the NEDI method, the proposed MEDI method will consider three more training windows centered at $Y_{2i,2j+2}$, $Y_{2i+2,2j}$ and $Y_{2i+2,2j+2}$, as illustrated by Figure 4(c), (d) and (e) respectively. The covariance signal energy of all training windows will be compared. The higher the energy in the training window, more likely the edge exists. The one contains the highest energy will be applied to the linear prediction in eq.(1). In this example, training window in Figure 4(c) is applied for the prediction. Similarly, the MEDI method considers six training window candidates in the second step, with such windows centered at $Y_{2i,2j-2}$, $Y_{2i,2j}$, $Y_{2i+2j+2}$, $Y_{2i+2,2j-2}$, $Y_{2i+2,2j}$ and $Y_{2i+2,2j+2}$ (see Figure 2 for

the pixel locations). The proposed method further adopts a hybrid framework, where the pixels at the edge region are interpolated by covariance-based method and the pixels at smooth region are interpolated by bilinear interpolation. If the variance of the pixels in the local block is larger than ε, the unknown pixel is regarded to be part of an edge, thus covariance-based method is applied.

4. RESULTS AND DISCUSSIONS

The proposed algorithm has been compared with other interpolation algorithms in literature including bilinear interpolation, the NEDI method [1], the IEDI method [3], the iNEDI method [4] and the ICBI method [6]. The proposed algorithm was implemented in Matlab running on a PC with Intel Pentinum(R) Duo Core 3 GHz CPU and 1GB DDR Ram. For comparison purpose, the IEDI method is also implemented in Matlab. Noted that heat diffusion refinement in the IEDI was bypassed because the investigation was mainly focused on the covariance mis-match problem. For bilinear interpolation, the built-in function in Matlab was used. For the rest of the methods, Matlab source code available on authors' websites were used [8–10]. The default function parameters of iNEDI and ICBI were applied. The threshold ε=48 was applied in the MEDI, NEDI and IEDI methods. The interpolation of the image boundaries was achieved by pixel extension. Both synthetic image and natural images were tested with different methods.

The original test image was first direct downsampled by a factor of two, that is from 2H×2W to H×W. The downsampled images were then expanded to their original sizes by using different interpolation methods. The interpolated images were compared with the original images objectively by measuring the PSNR and the *structural similarity index* (SSIM) [11]. The PSNR and *SSIM* of all test images are summarized in Table 1 and 2. The *PSNR* has been widely used to measure the distortion of the grayscale images after processing and given by

$$PSNR = 20log_{10}\left(\frac{255}{\sqrt{MSE}}\right), \quad (4)$$

$$MSE = \frac{1}{2H\times 2W}\sum_{i=0}^{2H-1}\sum_{j=0}^{2W-1} Z_{i,j}^2. \quad (5)$$

$$Z_{i,j} = |L_{i,j} - Y_{i,j}|. \quad (6)$$

where $L_{i,j}$ and $Y_{i,j}$ are the pixels in the original image and the interpolated image at location (i, j), respectively. Another objective measurement is the *SSIM*. The higher the *SSIM* value indicates there is a greater structural similarity between the original image and the interpolated image. The *PSNR* and *SSIM* of synthetic image "letter Y" are summarized in Table 1. The bilinear method shows the worst *PSNR* and *SSIM* values because of both the blurring effect and aliasing problem. The ICBI method shows the highest *PSNR* and *SSIM* values, because the synthetic image contains only high contrast edges, which are beneficial to isophate-based method. Among all the statistically optimal methods, the proposed method shows the highest *PSNR* and *SSIM* values because it does not only eliminate the error propagation problem, but also the covariance mis-match problem.

Besides the synthetic image, comparison has been performed using natural images, where the *PSNR* and *SSIM* results are summarized in Table 2. The bilinear method always shows the lowest *PSNR* and *SSIM* values. The *PSNR* and *SSIM* performance are observed to be image dependent, and there is no clear winner between the proposed method, the NEDI method, the iNEDI method and the ICBI method. Consider the image with less texture, like "grayscale F16", the ICBI method achieves the highest PSNR, and the iNEDI method achieves the second highest *PSNR*, then the proposed method and finally the NEDI method. However, for texture rich image, like "grayscale baboon", a reverse order is observed. Furthermore, the highest *PSNR* does not always imply the highest *SSIM* or vice versa. For example, consider the test image "boat". Although the NEDI method achieves the highest *PSNR*, it only achieves the fourth highest *SSIM* value. However, it is observed that the proposed method and the IEDI method demonstrate average *PSNR* and *SSIM* performance in a variety of images.

Method	*PSNR*	*SSIM*
MEDI	22.46	0.9352
Bilinear	19.75	0.8852
NEDI [1]	22.19	0.9311
IEDI [3]	22.26	0.9342
iNEDI [4]	22.19	0.9289
ICBI [6]	22.88	0.9456

Table 1: The *PSNR* and *SSIM* of test image "Letter Y" for different interpolation methods.

4.1 Subjective Test

Besides the objective measurement, subjective test was performed to evaluate the visual perception of the interpolated images. Error images (i.e. $Z_{i,j}$ in eq.(6)) are used as an evaluation tool. To obtain a fair comparison, the intensity of the error images are normalized with the same normalization factor among all interpolation method, and thus not all error images have their pixel values span from 0 to 255. Figure 5 shows the original image, interpolated images and the error images of test image "letter Y". It is observed that the MEDI interpolated image is perceptually more pleasant among all the interpolated images, because of the continue and smooth diagonal edges. It is more vivid by observing the error images. The white region concentrated along the edges in bilinear interpolated image because of the blurring problem. The white region is comparatively less obvious in the error images of the iNEDI and ICBI methods. The white region is dispersed in the NEDI case because the edges are interpolated by covariance matching, thus minimizing the error along edges. The white region is even more dispersed in the IEDI case, especially along the diagonal edges, because it fully utilizes the low resolution pixels with an enlarged training window. For the MEDI case, the white region is observed to be even dimmer and segmented along the diagonal edges, because the proposed method accurately adapts the edge orientation by covariance matching in multiple directions.

Figures 6 and 7 show the pixel intensity maps of the original image and the interpolated images of region A and region B in Figure 5, respectively. There is a sharp transition from 0 to 255 across the vertical edge of the original image in region A as shown in Figure 6. All the vertical edges are blurred after interpolation and the effect is the least significant for the iNEDI interpolated image, where the transition spanned 3 columns only. The blurring effect is the most vivid for the bilinear interpolated image. Halo effect is observed in the ICBI interpolated image. Similar interpolation performance are observed from the proposed method, the NEDI method and the IEDI method because these meth-

Image	Resolution Enhancement	MEDI		Bilinear		NEDI [1]		IEDI [3]		iNEDI [4]		ICBI [6]	
		PSNR	*SSIM*	*PSNR*	*SSIM*	*PSNR*	*SSIM*	*PSNR*	*SSIM*	*PSNR*	*SSIM*	*PSNR*	*SSIM*
Grayscale baboon	256×256⇒512×512	23.21	0.7123	22.27	0.6320	23.58	0.7361	23.24	0.7134	22.90	0.7280	22.47	0.7138
Bicycle	256×256⇒512×512	20.33	0.7790	18.56	0.6848	20.44	0.7726	20.41	0.7768	20.08	0.7812	18.90	0.7280
Boat	256×256⇒512×512	29.69	0.8911	27.06	0.8359	29.81	0.8910	29.75	0.8915	29.70	0.8920	29.21	0.8828
Grayscale F16	256×256⇒512×512	31.46	0.9326	28.34	0.8958	31.40	0.9308	31.46	0.9327	31.96	0.9375	32.44	0.9411

Table 2: The *PSNR* and *SSIM* of interpolated grayscale images by different interpolation methods.

ods using the same training window structure. Furthermore, the covariance structure is identical in all cases because it is a perfect vertical edge in the synthetic image. The outstanding performance of the proposed method is emphasized in the study of the intensity maps for region B as shown in Figure 7, which contains a diagonal edge. The interpolated edge of the bilinear method is the most blurred. Halo effect is also observed in the ICBI interpolated image. It is observed that the IEDI method achieves sharper diagonal edges than that of the NEDI method because a modified training window is applied in the second step of the IEDI method, which fully utilizes the original image information. The iNEDI method results in sharp and smooth edge but the edge continuity is not close to that of the original image. The proposed method forms sharp and smooth edge, the interpolated edge structure is highly close to the original edge, and thus shows the best subjective performance among different methods.

Figure 3 shows the simulation results for the test image "bicycle" with part of the original and interpolated images being zoomed-in. Consider the enclosed edges, the proposed method and the IEDI method show the most outstanding performance in preserving the continuity, smoothness and sharpness of the interpolated edge. The proposed method further preserves the image structure even at the edge termination (enclosed with rectangular box). This verifies that the proposed method is effective in eliminating the covariance mis-match problem. Therefore, though both the proposed method and the IEDI method show comparable objective performance, the proposed method outperforms IEDI in preserving image structure because of the use of multiple training windows.

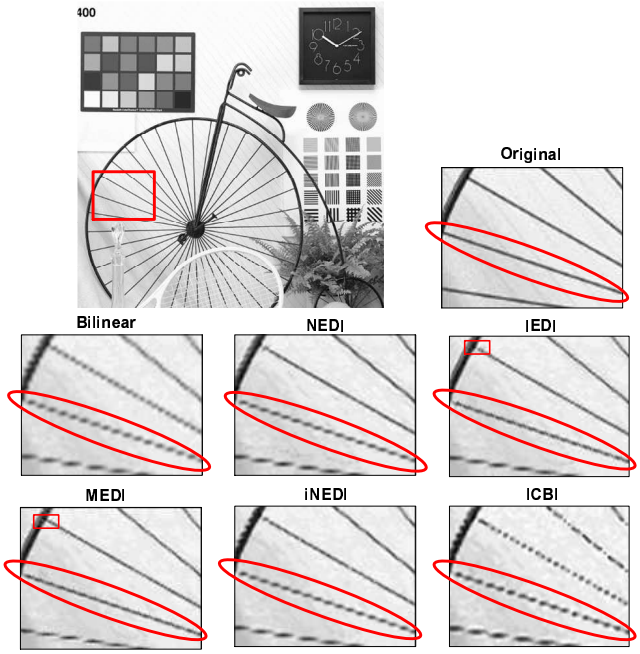

Figure 3: Original test image "bicycle" and zoomed-in portions of the original image and interpolated images.

5. CONCLUSION

An improved statistical optimized interpolation method, the *Modified Edge-Directed Interpolation* is presented. The proposed method overcomes the existing problems of new edge-directed interpolation by considering multiple training window and modified training window structure. The covariance mis-match problem is mitigated and the prediction error accumulation problem is eliminated. The performance of the proposed method has been verified with extensive simulation and comparison with other benchmark interpolation methods. Simulation results showed that the presented method has achieved outstanding perceptual performance with consistent objective performance that is independent to the image structure.

REFERENCES

[1] X.Li and M.T.Orchard, "New Edge-Directed Interpolation," *IEEE Trans. Image Processing*, vol.10, no.10, pp.1521-1527, Oct. 2001.

[2] M. Li and T.Q. Nguyen, "Markov random field model-based edge-directed image interpolation", *IEEE Trans. Image Processing*, vol.17, pp.1121-1128, July 2008.

[3] X.Q. Chen, J. Zhang, and L.N. Wu, "Improvement of a nonlinear image interpolation method based on heat diffuston equation," *Proc. Int. Conf. Machine Learning and Cybernetics*, pp.2911-2914, Nov. 2003.

[4] N. Asuni and A. Giachetti, "Accuracy improvements and artifacts removal in edge based image interpolation," *Proc. 3rd Int. Conf. Computer Vision Theory and Applications*, Madeira, Jan. 2008.

[5] M.J. Chen, C.H. Huang and W.L. Lee, "A fast edge-oriented algorithm for image interpolation," *Image and Vision Computing*, vol.23, pp.791-798, 2005.

[6] A. Giachetti and N. Asuni, "Fast Artifacts-Free Image Interpolation," *Proc. British Machine Vision Conference*, Leeds, Sept. 2008.

[7] W.S. Tam. "Modified edge-directed interpolation." M.Sc. thesis, Hong Kong Polytechnic University, Hong Kong, 2007.
Internet: http://library.polyu.edu.hk/record=b2080904∼S6.

[8] X. Li. "New edge-directed interpolation." Internet: http://www.csee.wvu.edu/∼xinl/source.html

[9] N. Asuni. "iNEDI (improved New Edge-Directed Interpolation)." Internet: http://www.mathworks.co.uk/matlabcentral/fileexchange/13470, Aug.6, 2007.

[10] A. Giachetti and N. Asuni. "ICBI download page." Internet: http://www.andreagiachetti.it/icbi/.

[11] Z. Wang, et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Processing*, vol.13, no.4, pp.600-612, April 2004.

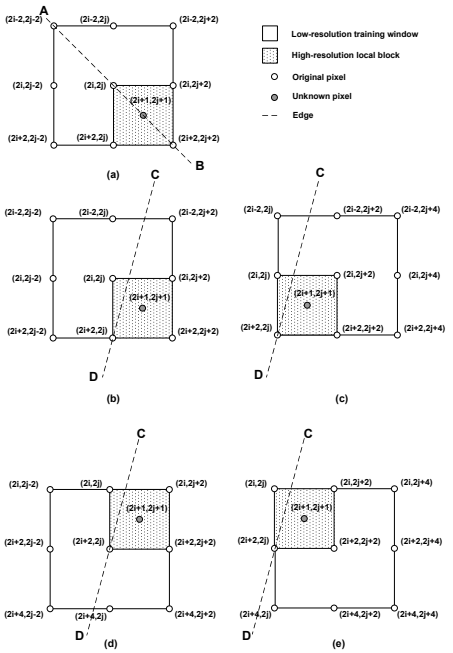

Figure 4: Illustration of (a) the single training window of the NEDI method with edge "AB" and (b–e) the four training windows of the MEDI method with edge "CD".

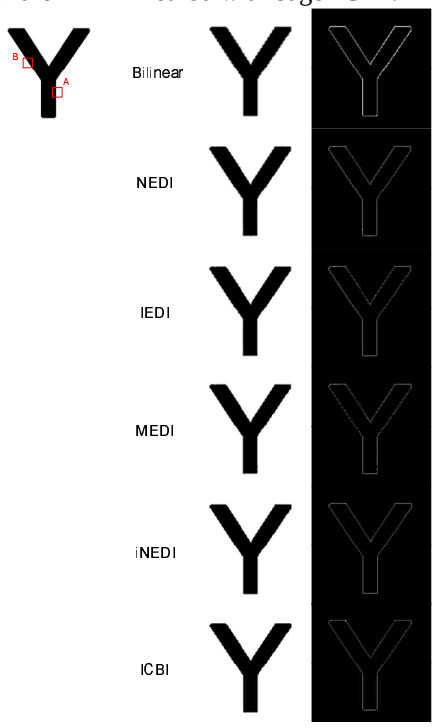

Figure 5: Original image, interpolated images and error images of "letter Y" (Resolution enhancement from 100×100 to 200×200).

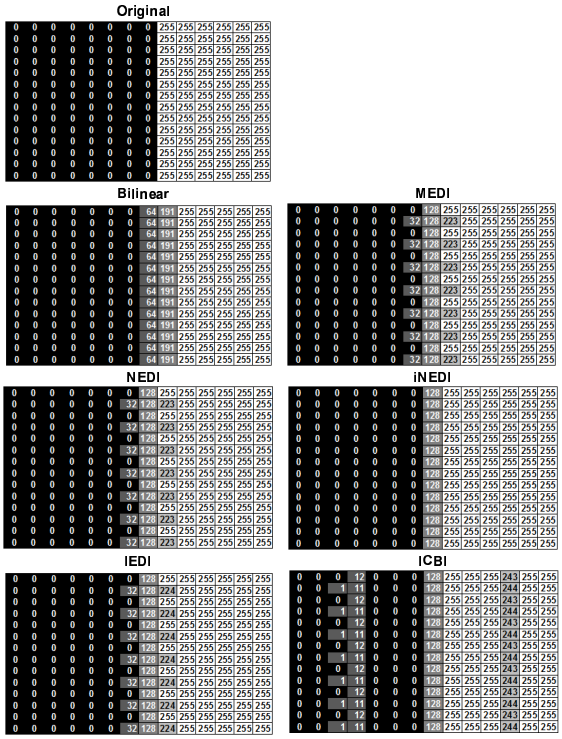

Figure 6: Pixel intensity maps of the original image and interpolated images of "letter Y" in region A.

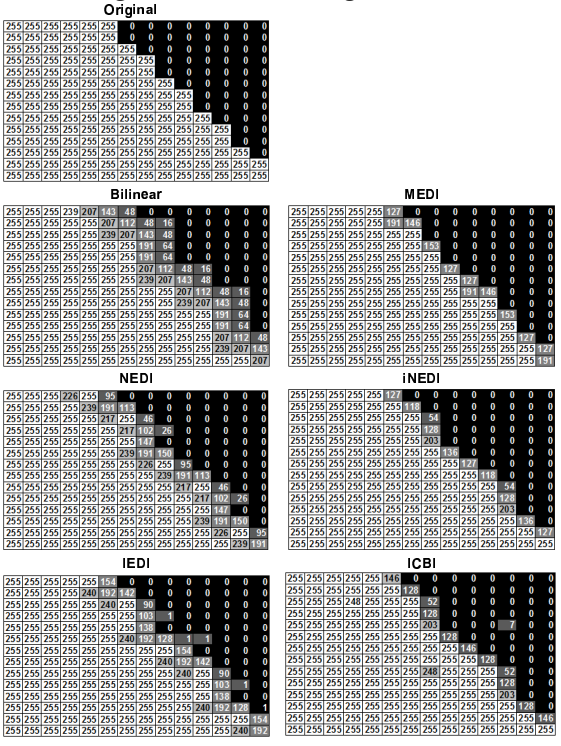

Figure 7: Pixel intensity maps of the original image and interpolated images of "letter Y" in region B.