Directive Cycle-Spinning Cellular Neural Network for High Frequency Subband Interpolation

Kosuke Hosaka†, Tsuyoshi Otake†, Hisashi Aomori‡,
Masatoshi Sato* and Mamoru Tanaka**

†Graduate School of Engineering, Tamagawa University,
6-1-1 Tamagawa Gakuen, Machida, Tokyo 194-8610, Japan
‡Department of Information System Technology, Chukyo University,
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan
*Faculty of System Design, Tokyo Metropolitan University,
6-6 Asahigaoka, Hiho, Tokyo 191-0065, Japan
**Department of Information and Communication Sciences, Sophia University,
7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan
Email: hskko0sc@engs.tamagawa.ac.jp, otake@ieee.org

Abstract– In this paper, we propose a novel image resolution enhancement technique based on the interpolation of high-frequency components of discrete wavelet transform (DWT). The low-resolution (LR) image whose resolution is to be enhanced is decomposed into four subbands by DWT on the basis of lifting method. Then only the high frequency subbands are interpolated with magnified factor α by using cycle spinning cellular neural network (CS-CNN). The CS-CNN is particularly well-suited for solving nonlinear optimization problems defined in space such as image processing tasks. In our algorithm, a directive architecture using a CS-CNN is developed to prevent the unnecessary smoothing of image detail. While a discrete-time cellular neural network (DT-CNN) transforms all high-frequency subbands of LR image into coefficients to predict the original subbands of high-resolution (HR) image using the A-template, the directive cycle spinning method is applied to estimate the optimal coefficients from individual outputs of the DT-CNN as above. Experimental results indicate that the proposed method produces better results than the conventional image resolution enhancement methods.

1. Introduction

Image resolution enhancement is a method to increase the number of pixels for generating a high-resolution (HR) image from its low-resolution (LR) version. It has been widely used in many image-processing applications such as super resolution, 4K TV, high-quality printing, medical imaging, and so on. Recently, image resolution enhancement methods in the wavelet domain have been discussed in many papers [1]-[3]. In wavelet based techniques, it is assumed that the LR image whose resolution is to be enhanced is a lowpass-filtered and decimated HR image. These methods are based on wavelet-domain zero padding (WZP) [1]. The WZP sets the LR image to a low-frequency subband of a discrete wavelet transform (DWT) and high-frequency subbands are composed of all-zero matrices, then the interpolated image is reconstructed by an inverse DWT (IDWT). Moreover, there are many algorithms to estimate the preserved high-frequency information from the given LR image. For example, Woo et al. utilize the statistical relationship between coefficients at lower level, and it is modeled by using a hidden Markov model [2]. Adaptive Learning for Zooming (ALZ) [3] proposed by Battiato et al. improves the details in the HR image by using the local gradient estimation.

In this paper, we propose a novel image resolution enhancement technique based on the architecture of a discrete-time cellular neural network (DT-CNN) with an arbitrary magnification parameter. The DT-CNN has been applied to many applications such as image compression, filtering, and recognition [4][5]. In our previous work, we showed that effective interpolation can be obtained using the DT-CNN with a cycle spinning (CS) architecture (CS-CNN). However, in the conventional CS architecture, applying all possible shifts within a local neighborhood causes the unnecessary smoothing of image detail especially around edges. In this work, we extended the CS-CNN prediction to prevent over-smoothing.

2. Basic Cycle Spinning Cellular Neural Network

2.1. Cycle Spinning Technique

Figure 1 shows a block diagram of the CS process. The CS considers a range of shifts and is set as

$$y = \mathrm{Ave} S_{-h}\left(T(S_h(x))\right), \quad h \in H \tag{1}$$

where x is an input signal, y is an output signal, S_h is a shift operator, S_{-h} is an unshift operator, and T is an analysis technique. H and h are the range of shifts and the shift value, respectively. First, the CS shifts the data, then it transforms the shifted data, and then it unshifts the transformed data. After repeating this for the range of shifts and averaging the results, denoised signal data is effectively obtained.

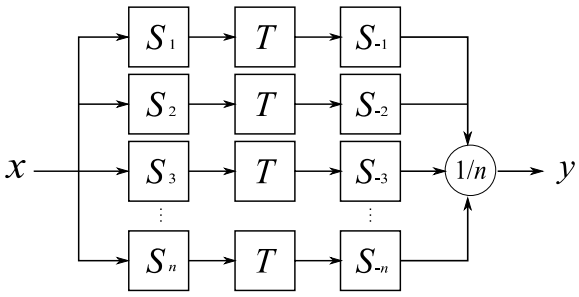


Figure 1: Cycle spinning process

2.2. DT-CNN

Figure 2 shows a block diagram of the DT-CNN. The state equation of the DT-CNN is described in matrix form as

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{f}(\mathbf{x}_n) + \mathbf{B}\mathbf{u} + \mathbf{T} \tag{2}$$

where **u** is an input vector, **x** is a state variable, **f**( ) is an output function, **A** and **B** are feedback and feedforward template matrices, and **T** is a constant vector, respectively. Let **y**=**f**(**x**), then the energy function E of the DT-CNN is defined as Lyapunov energy function.

$$E = -\frac{1}{2}\mathbf{y}^t(\mathbf{A} - \delta\mathbf{I})\mathbf{y} - \mathbf{y}^t\mathbf{B}\mathbf{u} - \mathbf{T}^t\mathbf{y} \tag{3}$$

$$\mathbf{D} = \delta\mathbf{I} = diag\{\delta, \cdots, \delta\} \tag{4}$$

where δ is a positive constant value used to determine the slope at a linear region of the output function. As described in [4] and [5], if the following conditions for the A-template are satisfied, it can be proved that the Lyapunov energy function E becomes a monotonically decreasing function.

$$A(i, j; k, l) = A(k, l; i, j) \tag{5}$$

$$A(i, j; k, l) > 0 \tag{6}$$

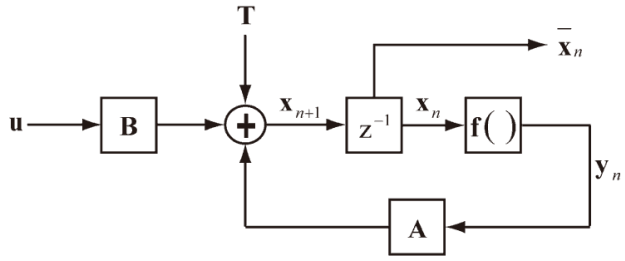


Figure 2: Discrete-time cellular neural network

2.3. CS-CNN

Figure 3 shows a block diagram of the CS-CNN. The state equation of the CS-CNN is described as

$$\mathbf{x}_{n+1} = \mathrm{Ave}\, S_{-d}\left(S_d(\mathbf{A})\mathbf{f}(\mathbf{x}_n) + \mathbf{B}\mathbf{u} + \mathbf{T}\right) \tag{7}$$

where S_d is a shift operator that applies horizontal and vertical shifts of (k, l), for example, (k-d, l), (k+d, l), (k, l), (k, l-d), and (k, l+d). d is the shift value of CS and S_{-d} is the unshift operator. This means that the center point of the A-template is shifted in each direction (vertical and horizontal) by the shift operator, and then, averaging over each unshifted convolution results in a shift in the A-template.

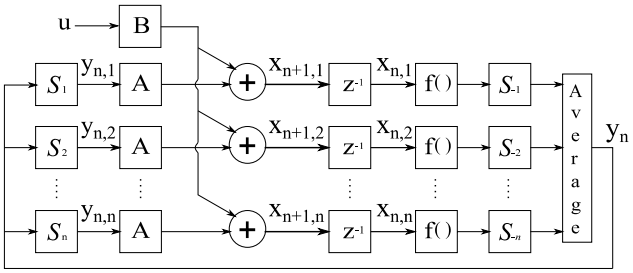


Figure 3: Cycle spinning cellular neural network

3. Subband Interpolation Using Directive CS-CNN

In the proposed method, subband images are interpolated using the two-layered CS-CNN with directive architecture. In the first layer of the CS-CNN, in order to obtain high prediction accuracy for subband images, it is necessary that the image can be reconstructed on the basis of the distortion function defined by

$$\mathrm{cost}(\mathbf{y}, \mathbf{u}) = \left\| \frac{1}{2}\mathbf{y}^t(\mathbf{G}\mathbf{y} - \mathbf{u}) \right\| \tag{8}$$

where ||.|| is the Euclidean norm and **G** is a Gaussian filter. This cost function means that not only the output ||**y**|| but also the difference between the interpolatively predicted subband image **Gy** and the input subband image **u** should be small. By the comparison between eqs. (3) and (8), the **A**-template, **B**-template, and threshold **T** can be determined as

$$\mathbf{A} = A(i, j; k, l), \quad C(k, l) \in N_r(i, j)$$

$$= \begin{cases} -(1+\lambda) & \text{if } k = i \text{ and } l = j \\ -\dfrac{1}{2\pi\sigma^2}\exp\left(-\dfrac{\{(k-i)^2 + (l-j)^2\}}{2\sigma^2}\right) & \\ \text{otherwise} & \end{cases} \tag{9}$$

$$\mathbf{B} = B(i, j; k, l), \quad C(k, l) \in N_r(i, j)$$

$$= \begin{cases} 0 & \text{if } k = i \text{ and } l = j \\ 1 & \text{otherwise} \end{cases} \tag{10}$$

$$\mathbf{T} = 0 \tag{11}$$

where σ is the standard deviation of the Gaussian function and λ is a regularization parameter. The B-template is only nonzero at the center value.

In our proposed method, directive architecture using a CS-CNN is developed. Figure 4 shows the shift patterns used for different directional CS-CNN. In order to prevent the over-smoothing, horizontal and vertical shift patterns are applied to wavelet decomposed LH and HL subband images, respectively. When the shift operator S_d is applied horizontally, such as (k-d, l), the A-template can be calculated by

$$S_d(\mathbf{A}) = S_d(A(i,j;k,l)) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(k-i-d)^2+(l-j)^2}{2\sigma^2}\right) \quad (12)$$

The other A-templates with different types of shift are calculate in the same manner. Then we can represent the dynamics of the first layer of the CS-CNN by using the above parameters as follows.

$$x_{ij}(t+1) = \mathrm{Ave}S_{-d}\left(\sum_{C(k,l)\in N_r(i,j)} S_d(A(i,j;k,l))y_{kl}(t) + u_{kl}\right) \quad (13)$$

$$y_{ij}(t+1) = f(x_{ij}(t+1)) \quad (14)$$

where $N_r(i, j)$ is the r-neighborhood of cell $C(i, j)$, expressed as $N_r(i, j)=\{ C(k, l) \mid \max\{ \mid k - i \mid, \mid l - j \mid \} \le \mathrm{r}\}$. $x_{ij}(t)$, $y_{ij}(t)$, and u_{ij} indicate the internal state, the output, and the input of cell $C(i, j)$, respectively. The output function $f(\ )$ corresponds to a piecewise linear function. The subband image is set to u_{ij} and the equilibrium output y^e_{ij} is obtained after the transition of the network.

Next, the output of the first layer of the CS-CNN becomes the input of the second layer of the CS-CNN, which has no dynamics, and the output of the second layer of the CS-CNN provides the predicted value. The pixel of the input image with coordinates (i, j) is mapped to the pixel (i', j') of the resolution enhanced image. Let d_m be an enlargement parameter, then the relationship between pixels (i, j) and (i', j') is determined as (i', j')=(id_m, jd_m). A deficient pixel of the enlarged image with coordinates (k, l) is obtained using

$$\hat{y}_{kl} = \mathrm{Ave}S_{-d}\left(\sum_{y_{kl}\in N(i',j')} S_d\left(\hat{B}(i',j';k,l)\right)y^e_{i'j'}\right) \quad (15)$$

The image enhancement is composed of horizontal process and vertical processes, as shown in figure 5. At the vertical resolution enhancement stage, we use the $\hat{\mathrm{B}}_v$-template, which is obtained by extending the A-template of the first layer of the CS-CNN vertically, that is,

$$\hat{\mathbf{B}}_v = \hat{B}_v(i',j';k,l), \quad C(k,l)\in N(i',j') = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(k-i)^2+(l-j')^2/{d_m}^2}{2\sigma^2}\right) \quad (16)$$

$$N(i',j') = \{C(k,l) \mid \max\{\mid k-i\mid, \mid l-j'\mid\} \le rd_m\} \quad (17)$$

In the same manner, at the horizontal resolution enhancement stage, we use the $\hat{\mathrm{B}}_h$-template, which is obtained using

$$\hat{\mathbf{B}}_h = \hat{B}_h(i',j';k,l), \quad C(k,l)\in N'(i',j') = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(k-i')^2/{d_m}^2+(l-j)^2}{2\sigma^2}\right) \quad (18)$$

$$N(i',j') = \{C(k,l) \mid \max\{\mid k-i'\mid, \mid l-j\mid\} \le rd_m\} \quad (19)$$

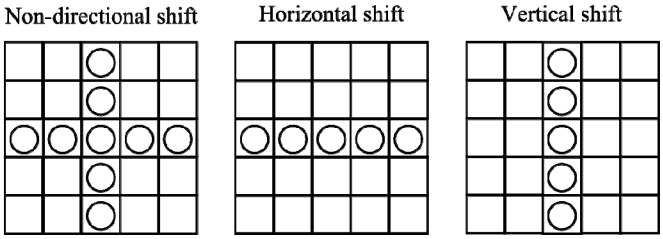
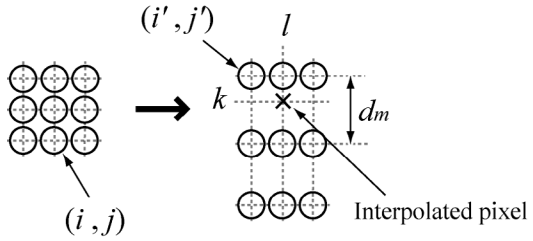
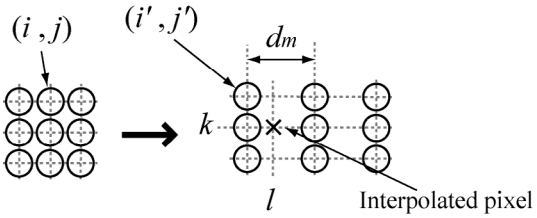


Figure 4: Directional shift patterns



(a) Vertical enlargement



(b) Horizontal enlargement

Figure 5: Image enlargement at each stage

4. Proposed Image Resolution Enhancement System

The deterioration of the resolution enhanced image is caused by the smoothing, that is, a general lack of high-frequency components. Accordingly, in order to improve the performance of resolution enhancement method, the prediction of the edges is absolutely imperative. Figure 6 shows the block diagram of the proposed image resolution enhancement system. When the input image size is $M\times N$, the subband image size becomes $M/2 \times N/2$. Here, we assume that the high frequency subbands include the survived lowpass-filtered edge information of HR image. Next, a directive CS-CNN is applied to high-frequency subbands with magnified factor α.

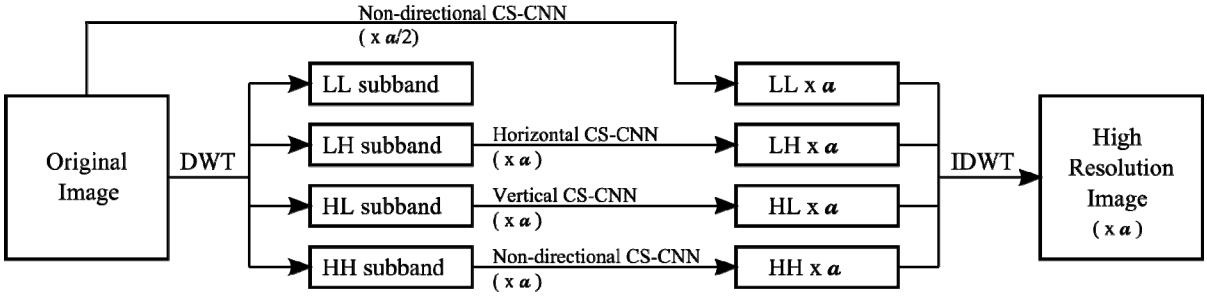


Figure 6: Proposed resolution enhancement system

As a result, we obtain the interpolated high frequency subbands (LHx, HLx, and HHx) with the size of $\alpha M/2 \times \alpha N/2$. These interpolated subbands are set to high frequency subbands of DWT for reconstruction of the HR image. Instead of the LL subband, which includes less high frequency information than the input LR image, we exploit the input LR image to predict the lowest frequency subband of DWT (LLx) for reconstruction. The input LR image is interpolated by CS-CNN with the half of the magnified factor $\alpha/2$, then, the interpolated LLx is set to the lowest frequency subband. Finally, HR image is obtained by applying the IDWT to the interpolated subbands, LLx, LHx, HLx, and HHx.

5. Experimental Results

In this section, we evaluate proposing novel image resolution enhancement algorithm using directive CS-CNN. We applied our system to the 8-bit gray-scale standard test images; Lena, Elaine, Sailboat, and Harbor. The HR version of these images with the size of 512×512 is used as the ideal interpolated image for performance evaluation purpose. These images are lowpass-filtered and downsampled to provide the LR images used for image enhancement.

The LR image is decomposed by lifting-based DWT using the well-known Le Gall 5/3 tap filter. The high frequency subbands are interpolated with magnified factor α and the LR image is also interpolated with magnified factor $\alpha/2$ independently by using directive CS-CNN interpolation. Then the HR image is obtained by applying the IDWT to interpolated subband images.

The performance of the proposed method was compared with that of the bilinear interpolation algorithm (BL), the bicubic interpolation algorithm (BC), and the conventional DT-CNN method (CNN). For the simulation, each parameter was decided experimentally; the standard deviation of the Gaussian was σ=0.6, the r-neighborhood of the cell was r=2, and the shift value of the cycle spinning was d=0.15. To enable a comparison with the resolution enhancement performance of WT, the enlargement parameter was set to α=2.

Table 1 shows the results for the peak signal-to-noise ratio (PSNR) values between the original images and enhanced images. Our results show that the proposed directive CS-CNN outperforms the conventional methods.

Table 1: PSNR (dB) results for 2× enlarged images (from 256×256 to 512×512).

Method/Image	Lena	Elaine	Sailboat	Harbor
BL	30.13	30.60	27.89	23.56
BC	31.34	31.17	29.18	24.01
CNN	32.15	31.82	29.25	24.12
Proposed	33.58	32.69	29.76	24.27

6. Conclusions

A new image resolution enhancement method operating in the wavelet domain was proposed. The main elements of this algorithm were the wavelet zero-padding based resolution enhancement and the modeling of the high frequency wavelet subbands by using the directive CS-CNN. The experimental results show that our proposed method consistently has better or competitive performances compared with conventional resolution enhancement methods.

References

[1] S. G. Chang et al., "Resolution enhancement of images using wavelet transform extrema extrapolation," *Proc. ICASSP95*, vol. 4, pp. 2379-2382, 1995.

[2] S. Zhao et al., "Wavelet Domain HMT-Based Image Superresolution," *Proc. ICIP03*, vol. 2, pp. 933-936, 2003.

[3] S. Battiato et al., "ALZ: Adaptive Learning for Zooming Digital Images," *Proc. ICCE07*, pp. 1-2, 2007.

[4] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, no. 10, pp. 1257-1272, 1988.

[5] N. Takahashi et al., "Nonlinear interpolative effect of feedback template for image processing by discrete-time cellular neural network," *J. Circuits Syst. Comp.*, vol. 12, no. 4, pp. 505-518, 2003.