Adaptive Cycle Spinning Cellular Neural Network for Image Resolution Enhancement

Kei Chiba¹, Tsuyoshi Otake¹, Hisashi Aomori²,
Nobuaki Takahashi³ and Mamoru Tanaka⁴

¹Graduate School of Engineering, Tamagawa University
6-1-1 Tamagawa Gakuen, Machida, Tokyo 194-8610, Japan

²Department of Information System Technology, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan

³IBM Japan, Ltd., 19-21 Nihonbashi, Hakozaiki-cho, Chuo-ku, Tokyo 103-8510, Japan

⁴Department of Information and Communication Sciences, Sophia University
7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554, Japan

E-mail: chbke8sc@engs.tamagawa.ac.jp, otake@ieee.org

Abstract

Cycle spinning cellular neural networks (CS-CNNs) are artificial neural networks that work effectively to solve large-scale problems. In our previous work, a CS-CNN is applied to enhance the resolution of images with an arbitrary magnification parameter. In this paper, a novel adaptive 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 pixel values into coefficients to predict the original pixel values using the A-template, the adaptive cycle spinning method with “minmod” functions is applied to estimate the optimal coefficients from individual outputs of the DT-CNN as above. The minmod functions are defined on the basis of the interpolation error theorem. 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. 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 low-resolution (LR) image whose resolution is to be enhanced is a lowpass-filtered and decimated high-resolution (HR) image. In other words, the LR image corresponds to the low-frequency coefficients of a wavelet-transformed 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 wavelet transform and high-frequency subbands are composed of all-zero matrices, then the interpolated image is reconstructed by an inverse wavelet transform. In [2], high-frequency coefficients are estimated using

coarser subband coefficients for high prediction accuracy. In [3], a hidden Markov model framework is introduced. However, due to the constraint of dyadic decomposition for the wavelet transform, the resolution enhancement factor is limited to 2^n ($n=1, 2, \dots$).

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]-[6]. Nonlinear interpolative dynamics using a feedback A-template is one of the significant characteristics of a CNN, enabling it to solve the optimization problem of minimizing the Lyapunov energy function. 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 match arbitrarily oriented edges in order 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

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

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

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

where $\mathbf{u}$ is an input vector, $\mathbf{x}$ is a state variable, $\mathbf{f}(\cdot)$ is a multilevel quantizing function, $\mathbf{A}$ and $\mathbf{B}$ are feedback and feedforward template matrices, and $\mathbf{T}$ is a constant vector, respectively. Let $\mathbf{y} = \mathbf{f}(\mathbf{x})$, then the Lyapunov energy function E of the DT-CNN is defined by

$$E = -\frac{1}{2} \mathbf{y}'(\mathbf{A} - \delta \mathbf{I})\mathbf{y} - \mathbf{y}'\mathbf{B}\mathbf{u} - \mathbf{T}'\mathbf{y} \quad (3)$$

$$\mathbf{D} = \delta \mathbf{I} = \text{diag}\{\delta, \dots, \delta\} \quad (4)$$

where δ is a positive constant value used to determine the quantizing region. As described in [5] and [6], 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.

$$A(i, j; k, l) = A(k, l; i, j) \quad (5)$$

$$A(i, j; k, l) > 0 \quad (6)$$

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

$$\mathbf{x}_{n+1} = \text{Ave } S_{-d}(S_d(\mathbf{A})\mathbf{f}(\mathbf{x}_n) + \mathbf{B}\mathbf{u} + \mathbf{T}) \quad (7)$$

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 $(k, l+d)$, as shown in Figure 4. 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

Shifted point

Figure 4: Shift operator of cycle spinning

3. Image Enhancement Using Two-Layered CS-CNN

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

$$\text{dist}(\mathbf{y}, \mathbf{u}) = \frac{1}{2} \mathbf{y}'(\mathbf{G}\mathbf{y} - \mathbf{u}) \quad (8)$$

where $\mathbf{G}$ is a Gaussian filter. This distortion function means that the difference between the interpolatively predicted image and the input image should be small. By the comparison between eqs. (3) and (8), the A-template, B-template, and threshold $\mathbf{T}$ can be determined as

$$\mathbf{A} = A(i, j; k, l), \quad C(k, l) \in N_r(i, j) \\ = \begin{cases} -(1 + \lambda) & \text{if } k = i \text{ and } l = j \\ -\frac{1}{2\pi\sigma^2} \exp\left(-\frac{\{(k-i)^2 + (l-j)^2\}}{2\sigma^2}\right) & \text{otherwise} \end{cases} \quad (9)$$

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

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

$$\mathbf{T} = 0 \quad (11)$$

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. When the shift operator S_d is applied horizontally, such as $(k-d, l)$, the A-template can be calculated by

$$S_d(\mathbf{A}) = S_d(A(i, j; k, l))$$

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

Then we can represent the dynamics of the first layer of the CS-CNN by using the above parameters as follows.

$$x_{ij}(t+1) =$$

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

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

where $N_r(i, j)$ is the r -neighborhood of cell $C(i, j)$, expressed as $N_r(i, j) = \{C(k, l) \mid \max\{|k-i|, |l-j|\} \leq r\}$. $x_{ij}(t)$, $y_{ij}(t)$, and u_{ij} indicate the internal state, the output, and the input of cell $C(i, j)$, respectively. The output function $f(\cdot)$ corresponds to the rounding operation. The output function plays an important role in optimizing the interpolation by considering the nonlinearity caused by quantization noises. The image pixels are set to u_{ij} and the equilibrium output y_{ij}^e 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

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

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 $\hat{B}_v$ -template, which is obtained by extending the A-template of the first layer of the CS-CNN vertically, that is,

$$\hat{\mathbf{B}}_v = \hat{B}_v(i', j'; k, l), \quad C(k, l) \in N(i', j')$$

$$= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(k-i')^2 + (l-j')^2/d_m^2}{2\sigma^2}\right) \quad (16)$$

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

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

$$\hat{\mathbf{B}}_h = \hat{B}_h(i', j'; k, l), \quad C(k, l) \in N'(i', j')$$

$$= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(k-i')^2/d_m^2 + (l-j')^2}{2\sigma^2}\right) \quad (18)$$

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

Figure 5: Image enlargement at each stage

4. Adaptive CS-CNN

Figure 6 shows a block diagram of our proposed adaptive CS-CNN, which introduces the concept of the interpolation error theorem. The state equation of the adaptive CS-CNN is described as

$$x_{ij}(t+1) =$$

$$\text{minmod}S_{-d}\left(\sum_{C(k,l) \in N_r(i,j)} S_d(A(i, j; k, l))y_{kl}(t) + u_{kl}\right) \quad (20)$$

Figure 6: Adaptive CS-CNN

$$\text{minmod } S_{-d}(\mathbf{y}) = \begin{cases} \text{Ave}(\mathbf{y}, \mathbf{y}_L, \mathbf{y}_R) & (\mathbf{y}_L - \mathbf{y})(\mathbf{y}_R - \mathbf{y}) > 0 \text{ and} \\ & \text{Ave}(|\mathbf{y}_L|, |\mathbf{y}_R|) < \text{Ave}(|\mathbf{y}_T|, |\mathbf{y}_B|) \\ \text{Ave}(\mathbf{y}, \mathbf{y}_T, \mathbf{y}_B) & (\mathbf{y}_T - \mathbf{y})(\mathbf{y}_B - \mathbf{y}) > 0 \text{ and} \\ & \text{Ave}(|\mathbf{y}_L|, |\mathbf{y}_R|) > \text{Ave}(|\mathbf{y}_T|, |\mathbf{y}_B|) \\ \mathbf{y} & \text{otherwise} \end{cases} \quad (21)$$

where y_L , y_R , y , y_T , and y_B indicate the output of each shifted CNN corresponding to the positions $(k-d, l)$, $(k+d, l)$, (k, l) , $(k, l-d)$, and $(k, l+d)$, respectively. The minmod function determines the direction of the error amendment.

5. Experimental Results

In this section, we evaluate the proposed novel image resolution enhancement algorithm using the adaptive CS-CNN. We applied our system to the 8-bit gray-scale standard test images, “Aerial,” “Airfield,” “Boat,” “Crowd,” “Lena,” and “Sailboat.” These were downsampled to provide the input LR images used for image enhancement.

The performance of the proposed method was compared with that of the bicubic interpolation algorithm (BC), the wavelet interpolation algorithm using the well-known Le Gall 5/3 tap filter (WT), and the conventional CS-CNN method (CS). For the simulation, each parameter was decided experimentally; the standard deviation of the Gaussian was $\sigma=0.6$, 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 $d_m=2$.

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 adaptive CS-CNN outperforms the conventional methods.

6. Conclusion

In the present work, a novel image resolution enhancement technique using a two-layered adaptive CS-CNN has been proposed. Our proposed method makes

good use of the nonlinear interpolative effect of the A-template and the error-amending ability of CS to obtain an enhanced-resolution image. The experimental results indicate that our proposed method consistently outperforms the conventional interpolation methods.

Table 1: Simulation results (from 256×256 to 512×512)

Image/Method	BC	WT	CS	Proposed
Aerial	26.501	26.839	27.086	28.102
Airfield	25.840	26.104	26.097	27.121
Boat	28.872	29.168	29.137	30.386
Crowd	32.673	32.104	32.969	33.786
Lena	33.750	33.473	33.978	35.012
Sailboat	27.891	29.109	29.140	30.468

References

- [1] S. G. Chang, Z. Cvetkovic and M. Vetterli: Resolution enhancement of images using wavelet transform extrema extrapolation, Proc. ICASSP95, Vol. 4, pp. 2379-2382, May 1995.
- [2] W. K. Carey, B. Chuang and S. S. Hemami: Regularity-preserving image interpolation, IEEE Trans. Image Process., Vol. 8, No. 9, pp. 1295-1297, 1999.
- [3] K. Kinebuchi, D. D. Muresan and T. W. Parks: Image interpolation using wavelet-based hidden Markov trees, Proc. ICASSP01, Vol. 3, pp. 7-11, May 2001.
- [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, T. Otake and M. Tanaka: 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.
- [6] H. Aomori, K. Kawakami, T. Otake, N. Takahashi, M. Yamauchi and M. Tanaka: Separable 2D lifting using discrete-time cellular neural networks for lossless image coding, IEICE Trans. Fundam., Vol. E88-A, No. 10, pp. 2607-2614, 2005.