2007 International Symposium on Nonlinear Theory and its Applications
NOLTA'07, Vancouver, Canada, September 16-19, 2007

Nonlinear prediction on image signals using radial basis function network

Keisuke Narisawa[†,1], Takashi Hoshino[†], Naoki Yabuta[†], and Tohru Ikeguchi[†]

†Graduate School of Science and Engineering, Saitama University
255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan
Email: [1]narisawa@nls.ics.saitama-u.ac.jp

Abstract—Differential pulse code modulation (DPCM) is one of the most popular methods to compress image signals. Although the DPCM generally works well, it could be flawed, because the DPCM predicts pixel values only by a linear function. Thus, the prediction errors might become large. In this paper, we proposed a nonlinear prediction method that incorporates the DPCM and a radial basis function (RBF) network. The proposed nonlinear prediction method reduces prediction errors of the DPCM by the RBF network. We confirmed that proposed method reduces prediction errors of the DPCM and average bit rates by numerical simulation to several standard images.

1. Introduction

Compression of image signals is inevitable to realize effective communication between end users. Differential pulse code modulation (DPCM) is one of the most popular methods to compress the image signals [1]. Because image signals generally have strong autocorrelation, the DPCM exhibits good performance. However, it has an intrinsic disadvantage: the DPCM only uses a linear function to predict pixels of an image, then prediction accuracy around edges of the image often decreases due to extreme change of the pixel values. To avoid such undesirable situation, several methods that improve prediction accuracy by adding nonlinear terms to the DPCM have been proposed [2, 3, 4]. Although these conventional methods often work well, they need a lot of computational complexity, because they use a supervised learning of a neural network. In this paper, we proposed a nonlinear prediction method with radial basis function (RBF) networks. Because the parameters of the RBF network can be decided using a least-square method, the computational complexity of the RBF network can be more reduced than that of a neural network with supervised learning. Thus, we proposed a nonlinear prediction method that combines the DPCM and the RBF network.

The proposed method has two essential parts. The first part is composed of the linear prediction using the DPCM applied to an image signal and the prediction errors of the DPCM are calculated. The second part is composed of a nonlinear prediction by the RBF. In the second part, the prediction errors of the DPCM are divided into three sets. This is a key point to predict prediction errors of the DPCM using the RBF network more efficiently. Next, nonlinear prediction using the RBF network is applied to prediction errors of the DPCM and the prediction errors of the RBF network are calculated. Finally, the prediction errors of the RBF network are encoded by the Huffman coding and average bit rates of the compressed image data are calculated. As a result, the proposed method shows high peak signal to noise ratio (PSNR) that quantifies prediction accuracy.

2. Differential pulse code modulation

The DPCM that predicts pixel values using their neighbor pixel values of horizontal and vertical directions is called two-dimensional (2-D) DPCM. Location of a predicted pixel p_0 and its neighbor pixels are shown in Fig.1. Here, a predicted value $\hat{p}_0$ of the pixel p_0 predicted by the 2-D DPCM is defined by

$$\hat{p}_0 = \sum_{m=1}^{M} a_m p_m \qquad (1)$$

where a_m $(m = 1, 2, \ldots, M)$ is the m-th coefficient, p_m is a pixel value, and M is the number of pixels used by the prediction.

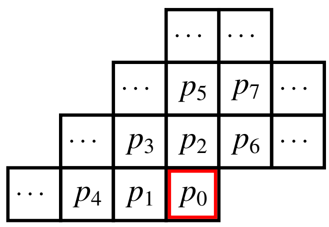


Figure 1: Pixel location of 2-D DPCM.

Then, the prediction error ϵ_0 by 2-D DPCM is defined as

$$\begin{aligned} \epsilon_0 &= p_0 - \hat{p}_0 \\ &= p_0 - \sum_{m=1}^{M} a_m p_m. \end{aligned} \qquad (2)$$

In the DPCM, these prediction errors are encoded. The coefficients are generally obtained by minimizing mean squared error of the prediction errors using least-square method.

3. Radial basis function network

In the RBF network, a nonlinear function is approximated by a sum of basis functions. The RBF network is defined by

$$\begin{aligned}\hat{y}_n &= \sum_{h=1}^{Q} c_h G(\boldsymbol{x}_n - \boldsymbol{\theta}_h) \\ &= \sum_{h=1}^{Q} c_h \exp(-\alpha_h |\boldsymbol{x}_n - \boldsymbol{\theta}_h|^2) \qquad (3)\end{aligned}$$

where $\hat{y}_n$ $(n = 1, 2, \ldots, N)$ is an output value, G is a basis function, and $\boldsymbol{x}_n$ is the D-dimensional input vector defined as

$$\boldsymbol{x}_n = (x_{n_1}, x_{n_2}, \ldots, x_{n_D}) \qquad (4)$$

where $x_{n_1}, x_{n_2}, \ldots,$ and x_{n_D} are pixel values used to approximate y_n. In this paper, Gaussian function is used as a basis function. Q is the number of the basis functions that approximate the nonlinear function, c_h is the height of the basis function, α_h is the spread of the basis function, $\boldsymbol{\theta}_h$ is the central coordinate of the basis function. These parameters can be calculated by the least-square method. A schematic example of approximating a nonlinear function form by the RBF network is shown in Fig.2 in case that $D = 1$, and $Q = 2$.

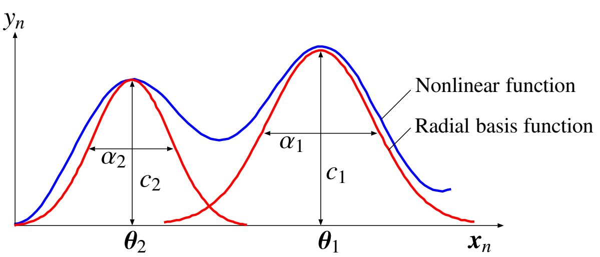


Figure 2: How to approximate a nonlinear function (the blue curve) by the RBF networks (red curves).

4. Proposed method

In this section, we explain the proposed method. The proposed method reduces prediction errors of the DPCM by the RBF network. The flow chart of the proposed method is shown in Fig.3 where P.E. means the prediction errors.

In the proposed method, at first, the linear prediction by the DPCM is applied to the image data and the prediction errors of the DPCM are calculated. Prediction error between predicted value $\hat{p}_0$ and the true value p_0 is defined by Eq.(2).

Next, input vectors applied to the RBF network are composed by an arbitrary combination of the prediction errors of the DPCM. Location of the prediction errors of the DPCM are shown in Fig.4. For example, an M-dimensional input vector $\boldsymbol{x}_0$ applied to the RBF network to predict ϵ_0 is

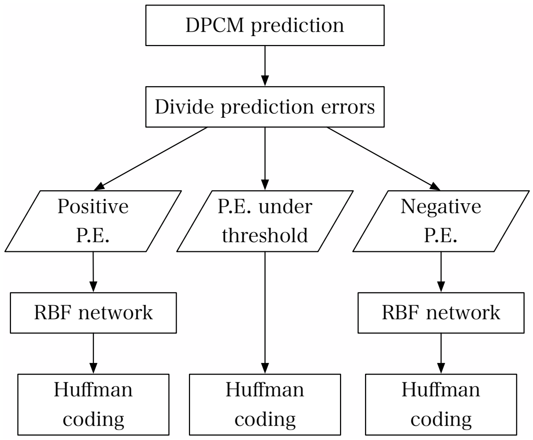


Figure 3: Flow chart of the proposed method.

$$\boldsymbol{x}_0 = (\epsilon_1, \epsilon_2, \ldots, \epsilon_M). \qquad (5)$$

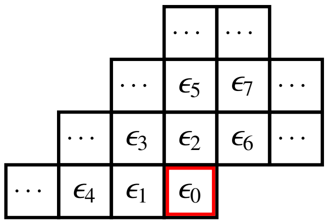


Figure 4: Location of the prediction errors by the DPCM.

Then, the prediction errors of the DPCM are divided into three groups with a threshold value T as follows;

1. $\epsilon_0 > 0, |\epsilon_0| > T$
2. $\epsilon_0 < 0, |\epsilon_0| > T$
3. $|\epsilon_0| \leq T$

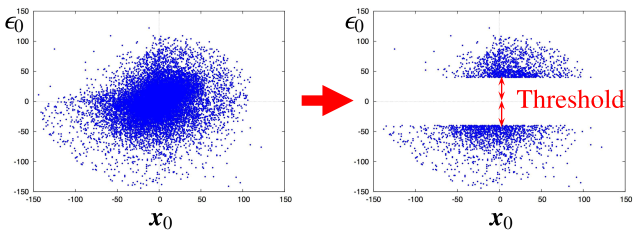


Figure 5: An example of dividing the prediction errors of the DPCM.

An example of this procedure is shown in Fig.5. The procedure described above is a process to apply a nonlinear prediction method using RBF networks to the prediction errors whose absolute values are large. It means that the prediction efficiency using the RBF network becomes high.

Next, the nonlinear prediction by the RBF network is applied to the positive and negative prediction errors of the DPCM whose absolute values are larger than the threshold T. Then, the prediction errors of the RBF network are calculated. The prediction error η_0 of the RBF network is defined as

$$\eta_0 = \epsilon_0 - \sum_{h=1}^{Q} c_h \exp(-\alpha_h |\boldsymbol{x}_0 - \boldsymbol{\theta}_h|^2) \qquad (6)$$

where Q is the number of the basis functions, c_h is the height of the basis function, α_h is the spread of the basis function, $\boldsymbol{\theta}_h$ is the position of the h-th radial basis function. Finally, the prediction errors of the RBF network are encoded by Huffman coding. Then, average bit rate is calculated.

Using this proposed method, the dynamic range of the prediction errors is much reduced. The image data is restored by these prediction errors and parameters of the DPCM and RBF network.

5. Experimental conditions

To evaluate performance of the proposed method, we used standard image data base (SIDBA). In the experiment, we use three standard images from SIDBA: Cameraman, Lenna, and Lighthouse. The size of these images is 256×256 pixels and they have 8-bits resolution. Prediction accuracy is estimated by the peak signal to noise ratio (PSNR) S defined as

$$S = 10 \log_{10} \frac{255^2}{\sigma_\epsilon^2} \quad [\mathrm{dB}] \qquad (7)$$

where σ_ϵ^2 is the mean square error of the RBF network.

The number of the basis functions in the RBF network, Q, is set to five, considering a good balance between the prediction performance and CPU time of calculation. Relation between the number of the basis functions and CPU time for Lenna is shown in Fig.6. When the number of the basis function is set to five, high performance of the prediction can be obtained by short CPU time. Then, we fix the number of RBF functions to five. As the pixel values used by the DPCM and the prediction errors used by the RBF network, all combinations of seven values in Figs.1 and 4 are used.

6. Experimental results

The results are shown in Table 1. Three results are the best performance among all combinations of pixel values used by the DPCM and the RBF network.

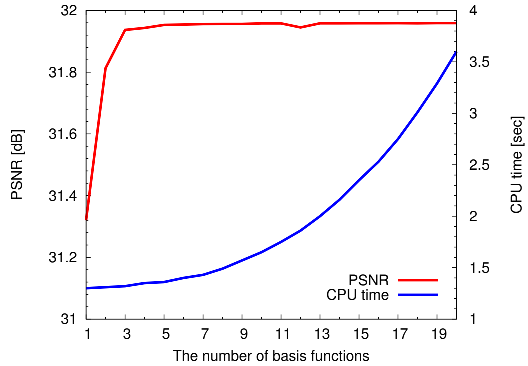


Figure 6: Relation between the number of the basis functions and CPU time for Lenna. The left vertical axis shows S [db], and the right vertical axis shows CPU time.

Table 1: Results of PSNR.

Image	PSNR [dB] 2-D DPCM	PSNR [dB] Proposed method
Lenna	28.3	34.2
Cameraman	24.9	30.2
Lighthouse	22.8	29.0

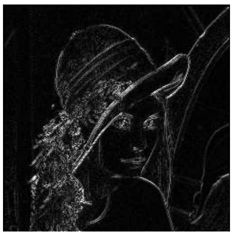
Table 2: Results when average bit rate is minimum.

Image	PSNR [dB] (2-D DPCM)	Bit rate [bits/pel] (2-D DPCM)	PSNR [dB] (Proposed method)	Bit rate [bits/pel] (Proposed method)
Lenna	28.3	4.43	31.9	4.41
Cameraman	24.9	4.68	25.8	4.68
Lighthouse	22.8	4.85	23.0	5.46

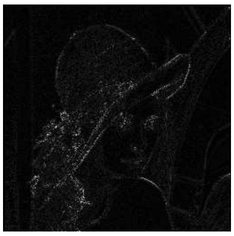
From Table 1, the prediction accuracy using the RBF network is higher than that of the DPCM. The original image of Lenna is shown in Fig.7(a). An image representation of prediction errors only by the DPCM is shown in Fig.7(b). The brighter pixels indicate that the prediction errors are large. The prediction error image obtained by the proposed method is shown in Fig.7(c). The prediction errors of the proposed method are reduced more than Fig.7(b).

Next, Table 2 shows the results when each average bit rate becomes minimum for all combinations of the pixel values by the DPCM and the RBF network. From Table 2, when the proposed method is applied to Lenna, it is seen that the prediction accuracy is improved and the average bit rate is reduced compared to that of the DPCM. Prediction error image by the proposed method is shown in Fig.7(d). When the proposed method is applied to Cameraman, it is seen that the prediction accuracy is improved, but the average bit rate exhibits almost the same performance as the DPCM. When the proposed method is applied to Lighthouse, the prediction accuracy is improved slightly, but the average bit rate is increased compared to that of the DPCM. Because the proposed method divides the prediction errors of the DPCM into three groups, the amount of information to encode may increase. In particular, the performance of the nonlinear prediction is not very good when the proposed method is applied to Lighthouse.

(a) Original



(b) Error image by the DPCM



(c) Error image by the proposed method when PSNR is maximum

(d) Error image by the proposed method when the average bit rate is minimum

Figure 7: Results by the proposed method (Lenna). Here, the values of the pixels are set to four times prediction errors to confirm the prediction errors easily.

To compare the performance of the proposed method with those of the conventional methods [2, 4], the experimental condition of the proposed method is set to same as the conventional methods.

We compared the performance to the conventional method by J. Li et al. [2], and the proposed method are shown in Table 3. To compare the performance fairly, the one-dimensional DPCM using the pixel values of only horizontal direction is used as a linear prediction, the number of the pixel values used by the DPCM is one, and the number of the pixel values used by a nonlinear prediction is four. In Table 3, it is seen that the prediction accuracy of the proposed method is higher than that of the conventional method [2].

We compared the results by S. A. Dianat et al. [4] (Table 4). In this comparison, the 2-D DPCM is used as a linear prediction, the number of the pixel values used by the DPCM and a nonlinear prediction is three. In Table 4, it is seen that the prediction accuracy of the proposed method is again higher than that of the conventional method [4].

Table 3: Comparison of the conventional method by J. Li et al. [2], and the proposed method.

Image	PSNR [dB]	
	Conventional method	Proposed method
Lenna	27.22	29.09

Table 4: Comparison of the conventional method by S. A. Dianat et al. [4], and the proposed method.

Image	PSNR [dB]	
	Conventional method	Proposed method
Lenna	31.20	32.01

7. Conclusions

In this paper, we proposed a nonlinear prediction method that combines the DPCM and the RBF network. We applied the proposed method to image data from SIDBA and calculated prediction accuracy and average bit rate. As a result, we confirmed that the proposed method improves the prediction accuracy and average bit rates compared to the conventional methods, however, we also confirmed a result that the average bit rates increased. The research of TI is partially supported by Grant-in-Aid for Scientific Research (B) from JSPS (No.17500136).

References

[1] N. S. Jayant, "Digital Coding of Speech Waveforms PCM, DPCM, and DM Quantizers," *Proc. IEEE*, Vol.62, No.5, pp.611–632, 1974.

[2] J. Li et al., "Nonlinear Prediction in Image Coding with DPCM," *Electronics Letters*, Vol.26, No.17, pp.1357–1359, 1990.

[3] Z. He et al., "Nonlinear Predictive Image Coding with a Neural Network," *Proc. ICASSP*, pp.1009–1012, 1990.

[4] S. A. Dianat et al., "A Non-linear Predictor for Differential Pulse-code Encoder (DPCM) using Artificial Neural Networks," *Proc. ICASSP*, pp.2793–2796, 1991.