Application of Neual 'Gas' Model in Image Compression

Bai-ling Zhang [1,*] Min-yue Fu[1] and Hong Yan [2]
[1]Department of Electrical and Computer Engineering
The University of Newcastle, Callaghan, NSW 2308, Australia
[2]Department of Electrical Engineering
The University of Sydney, NSW, 2006, Australia

Abstract *Topologically ordered vector quantization has attracted attention in recent years. In this paper, we propose to apply a topology representing learning algorithm — neural 'gas' model for obtaining topology ordered codebook for the vector quantization (VQ) and exploit it on image compression. Compared with the well-known Kohonen's self-organizing map (SOM), neural 'gas' model has several advantages, including faster convergence and higher signal-to-noise ratio in reconstruction. We illustrate some experimental results and discuss several relevent research issues.*

1. Introduction

Image compression is an essential task for image storage and transmission. Among various image compression techniques, vector quantization techniques may be preferable when higher compression rates are required. A vector quantizer (VQ) is a mapping of input vectors into a finite collection of codevectors and the problem of the design of a VQ consists in finding a codebook which minimizes the quantization mean square error.

In the neural network field, some competitive learning algorithms can be considered as on-line versions of the traditional k-means clustering or LBG algorithm for VQ. An important biologically inspired model self-organizing map (SOM), which forms a low dimensional structured representation of input data, has also been successfully applied to VQ [4-5]. Topologically ordered codebook has been proven to be able to bring at least the following advantages[4-5]: (1) the coding/decoding process is more robust with respect to channel errors; (2) a reduced image bit rate and a faster search phase are available; (3) it provides a possible progressive transmission (PT) procedure in communications systems.

*email: bailing@ee.usyd.edu.au

Kohonen's feature map is a special way for conserving the topological relationships in input data, which also bears some limitaitons. In SOM, the neighborhood relations between neurons must be predefined. More generally, the neighborhood relations between neurons should match the topological structure of the data manifold. On the other hand, SOM does not minimizes a global cost function and therefore is not optimal in approximating the input distribution. In recent years, developments on SOM have provided some alternatives. A typical example is the topology representing network called neural 'gas' model[(2-3)], which has a keypoint of introducing a neighborhood-ranking within the input space. To find the neighborhood-rank, each neuron compares its distance with the input vector to the distances of all the other neurons. Unlike SOM which requires a prior defined static neighborhood relations, neural gas model determines a dynamical neighborhood relations as learning proceeds. Neural 'gas' model has been successfully applied to some pattern recognition tasks, but its practical applications are still few comparing with the famous SOM.

The main purpose of this paper is to illustrate the advantages of the neural 'gas' model over SOM in constructing a good codebook for image compression. The exploitation of the topology preservation property will not be discussed. The faster convergence and smaller distortion errors, as pointed out in [2,3] for time-series prediction and further confirmed by our experiment on image coding, suggest the practical significance of neural 'gas' model in real data coding/decoding system.

This paper is organized as follows. In the next section, we briefly outline the principles of SOM and neural 'gas' model, and the ways of applying them in constructing codebook in vector quantization. Section 3 offers specific image compression experiments with SOM and neural 'gas' algorithm, respectively.

0-7803-4859-1/98 $10.00©1998 IEEE

Concluding remarks are provided in last section.

2. Self-organizing Map and Neural 'Gas' Model for Vector Quantization

Vector quantization techniques encode a data manifold V using a finite set of reference vectors (codebook). A reference vector $\mathbf{w}_k$ is considered to best match a data vector $\mathbf{x}$ in the sense that appropriately defined distortion measure such as the squared error $\|\mathbf{x}-\mathbf{w}_k\|^2$ is minimal. The codebook should partition the input manifold into the so called Voronoi polygons defined as

$$V_k = \{\mathbf{x} \in \mathrm{V} | \; \|\mathbf{x}-\mathbf{w}_k\| \leq \|\mathbf{x}-\mathbf{w}_l\|, \forall l\} \quad (1)$$

The problem can be described as an optimization issue with the cost defined as

$$J = \int dx P(\mathbf{x}) \|\mathbf{x}-\mathbf{w}_k\|^2 \quad (2)$$

where $P(\mathbf{x})$ is the distribution of input data. The traditional k-means clustering is the most straightforward way to minimizing (2) via gradient descent on E, which bears such disadvantages as local minima and underutilization of reference vectors, etc.

Kohonen's self-organizing map is neurobiology motivated neural learning algorithm, which can be applied to vector quantization. In SOM, every reference vector $\mathbf{w}_k$ is attached to a location k of a lattice. When presenting a data vector $\mathbf{x}$, not only the best matching reference vector $\mathbf{w}_k$ is adjusted, but also those $\mathbf{w}_l$ with l adjacent to k are updated, with a step size decreasing with the lattice distance between k and l. The learning algorithm is of the form

$$\Delta \mathbf{w}_k(\mathrm{m}) = \mu_k h_\sigma(\mathrm{k}, l)(\mathbf{x}-\mathbf{w}_k(\mathrm{m})), \; \mathrm{m} = 1, \cdots, \mathrm{M} \quad (3)$$

where h_σ is the neighborhood interaction function that decreases monotonically with increasing distance between k and l, which is typically taken as a Gaussian function $h_\sigma(\mathrm{k}, l) = \exp(-\frac{\mathrm{d}_{\mathrm{k}l}^2}{2\sigma^2})$.

In the neural gas model, the reference vectors $\mathbf{w}_m$ are adapted by the relative distance between the neural units within the input space, not by a topological arrangement of the neural units within a prestructured lattice. Each time an input $\mathbf{x}$ is presented, we first make an ordering of the elements of a set of distortions $E_{\mathbf{x}} = \{\|\mathbf{x}-\mathbf{w}(\mathrm{m})\|, \mathrm{m} = 1, \cdots, \mathrm{M}\}$ and then determine the adjustment of reference vector $\mathbf{w}(\mathrm{m})$. The neighborhood relation is described by a dynamic graphic structure called Delaunary triangular (DT). As the weight vectors change, the connectivity of the underlying DT graph changes dynamically. A connectivity matrix C is defined to represent the graph structure, with its elements $C_{ij} \geq 0$.

For a given data vector $\mathbf{x}$, we determine a "neighborhood-ranking" $(E_{\mathbf{x}}(m_0), E_{\mathbf{x}}(m_1), \cdots, E_{\mathbf{x}}(m_{M-1}))$ of the distortion set, with $\mathbf{w}_{m_0}$ being closest to $\mathbf{x}$, $\mathbf{w}_{m_1}$ being second closest to $\mathbf{x}$, $\mathbf{w}_{m_k}$, $\mathrm{k} = 0, \cdots, \mathrm{M}-1$ being the reference vector for which there are k vectors $\mathbf{w}_j$ with $\|\mathbf{x}-\mathbf{w}_j\| < \|\mathbf{x}-\mathbf{w}_{m_k}\|$. The creation of a "virtual" connection between m_0 and m_1 is described by setting $C_{m_0 m_1}$ from zero to one. At the same time, each neuron adjusts its own weight via dynamical learning rate which depends on the ranking of its representation capabilities. Denote the number k associated with each neural unit m by k_m. The following learning rule is the simple neural 'gas' algorithm in[2-3].

$$\Delta \mathbf{w}_k(\mathrm{m}) = \mu_k h_\lambda(\mathrm{k_m})(\mathbf{x}-\mathbf{w}_k(\mathrm{m})), \; \mathrm{m} = 1, \cdots, \mathrm{M} \quad (4)$$

where $h_\lambda(k_m)$ is 1 for $k_m = 0$ and decays to zero for increasing k_m with characteristic decay constant. In the simulaiton we choose the same one as in [3], $h_\lambda(k_m) = exp(-k_m/\lambda)$.

Each connection from node i to node j has an age t_{ij} that is the number of adaptation steps t the connection already exists without having been refreshed. If the age of a connection exceeds its lifetime T, the connection is removed. Connections have to die out because weight vectors that are neighboring at an early stage might not be neighboring any more at a more advance stage. However, a connection does not die out if it is regualrly refreshed.

3. Image Compression Experiments

A VQ based image compression system begins by mapping the image into a set of vectors, some of which are chosen as a training set. Next a library of reference vectors is generated using an appropriate VQ algorithm. Then each input vector is subsequently compared to the codewords in the codebook and is coded by the closest one, according to the distortion measure used. The label of the closest codeword in the codebook is transmitted in lieu of the input vector, thus achieving compression. The reconstruction of the image is accomplished by simple table look up. The receiver has an exact copy of the codebook with associated lables for each codeword. As the label for each vector is received, it is replaced by the corresponding reproduction vector in the codebook.

The first image used for the simulation is 256×256 pixels "Lena" image on a grey scale of 8 bits. The image was partitioned into a number of blocks of size 4×4 which were then arranged into series of one-dimensional input vectors of size 16. The mean vector $\mathbf{m}$ is calculated beforehand and stored. This mean is substracted from all randomly chosen subimages in the training sequence. The data at input layer is scaled from the original pixel values to continuous value (0-1). 256 codewords were used, which mean 0.5bpp compressed bit rate according the following definition

$$R = \frac{[\log_2 D]}{m} \tag{5}$$

where D is the number of codebook entries and m is the dimension of training vectors. When training with SOM algorithm, the 256 output units are arranged in 16×16 lattice and the width σ of neighborhood interaction function is initially large (5) and then shrink to small value (0.5). When experimenting with the neural gas model, the decay constant λ is dynamically changed from 30 to 0.005. Both algorithms are proceeded in 20 cycles, within each cycle 4096 randomly chosen samples participate in learning. As to the quality of reconstructed images, the peak signal-to-noise ratio (PSNR)

$$\mathrm{PSNR} = 10 log_{10} \frac{255^2}{\mathrm{MSE}}$$

is used as the performance measure, where the MSE denotes the mean square error over the entirety. Figure 1 gives a compression result. The original "Lena" image and the corresponding reconstructed images from SOM-VQ and Gas-VQ are shown in left column. The comparison of PSNRs is illustrated in Figure 2 (a).

Another image we used is a 480×500 mandrill image. To save the computing time, we rescale the image to a 240×250 with 8 bits gray scale. The codebook with 256 codewords is constructed in the same way as for Lena image. The compressed bits rate is 0.4 in this situation. SOM and the neural 'gas' models were also compared with the reconstructed images shown in the right column in Figure 1 and PSNR illustrated in Figure 2 (b). It is obvious that the codebook from neural 'gas' model provide much better reconstruction quality.

4. Discussions and Conclusions

The SOM based codebook could present a topological order between the codewords, which is a useful property especially when an efficient code search is needed. In this paper, we showed that the neural 'gas' algorithm as a topology representing model can be a beneficial alternative in constructing topologically ordered codebook for vector quantization and corresponding image compression. The fast convergence property and high reconstruction quality will find more significance in establishing a real vector quantization system. Currently we are further improving the neural 'gas' based VQ by combining some other techniques such as hierarchical VQ, VQ with edges classifiers, etc.

References

[1] T.Kohonen, "The self-organizing map," *Proc. IEEE* 78, pp.1464-1480, 1990

[2] T.M.Martinetz, S.G.Berkovich and K.J.Schulten, " 'Neural-gas' network for vector quantization and its application to time-series prediction," *IEEE Trans. Neural Networks* 4, pp.558-568, 1993.

[3] T.M.Martinetz, K.J.Schulten, "A 'neural-gas' network learns topologies," in *Artificial Neural Networks* T.Kohonen et al eds., Elsevier Science Publishers, 397-402, 1991

[4] G.Poggi, "Applications of the Kohonen algorithm in vector quantization," *European Transactions on Telecommunications and Related Techniques,* 6(2), pp.191-202, 1995

[5] S.Carrato, "Image vector quantization using ordered codebooks: properties and applications," Signal Processing, 40, pp.87-103, 1994

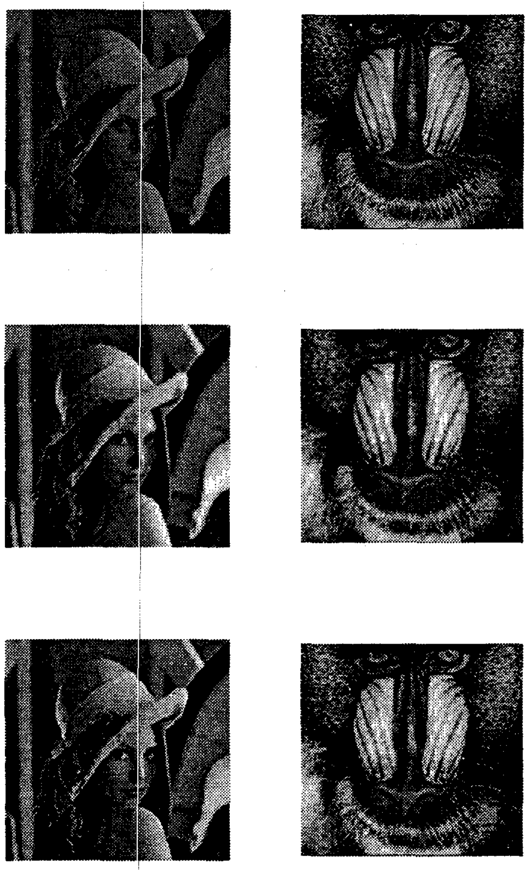

Figure 1: Left column: original Lena image (top), reconstructed image by 256 codewords from SOM-VQ (middle) and reconstructed image from Gas-VQ, also using 256 codewords (bottom). Right column: original mandrill image (top), reconstructed image by 256 codewords from SOM-VQ (middle) and reconstructed image by 256 codewords from Gas-VQ (bottom).

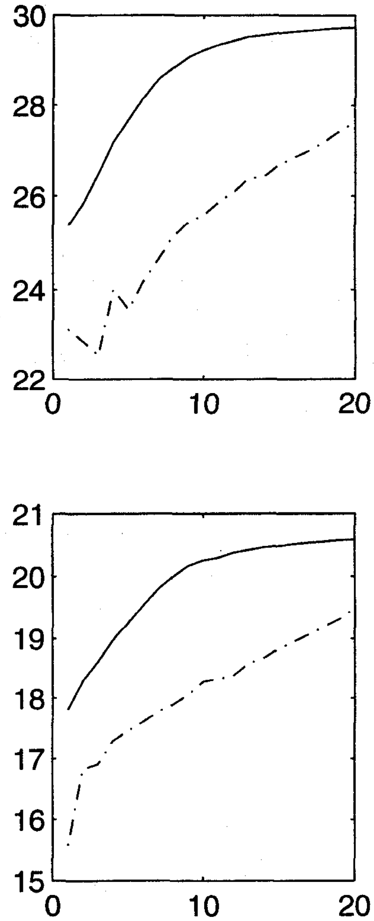

Figure 2: The peak signal-to-noise ratios (PSNR) vs training cycles for the reconstructed images from SOM (dot-and-dash line) and neural 'gas' model (solid line), respectively. (top) PSNR for reconstructed Lena image. In each cycle, randomly chosen 4096 4×4 blocks participate in training ; (bottom) PSNR for reconstructed mandrill image. In each cycle, randomly chosen 3000 4×5 blocks are used in training.