

A LEARNING VECTOR QUANTIZATION FOR RECOGNITION OF INVARIANT SPATIAL DETECTORS OF PERCEPTUAL PATTERNS CONSTITUENTS

Hanan Hassan Ali Adian^{1*}, Abd. Rahman Ramli², Bachok Taib³, and Adznan Jantan¹

Received: May 7, 2003; Revised: Jan 12, 2004; Accepted: Jan 29, 2004

Abstract

This paper investigates the use of a learning vector quantization (LVQ) neural network trained by invariants and spatial detectors for pattern recognition. Satellite imagery patterns are sensitive to translation, rotation, and scale variability. This motivates the construction of such detectors to constitute perceptual instances for an LVQ network. Influence of certain factors such as the hidden layer neurons, and the learning parameter are investigated.

Keywords: LVQ, invariant recognition, spatial characteristics, satellite imagery

Introduction

Feedforward neural networks turned to be attractive trainable machines for feature-based recognition. However, some problems arise when the domain of application is coarse and uneasy to discern between classes. This problem arises in the field of remote sensing because the domain of applicability inherits many variations. One of the complicated variations is the type and nature of the data obtained from satellite imagery. To end with a robust recognizer for such application is still a matter of experiments. This opens a wide area of research, especially with the vast existing methods. The heart of the problem can be solved by constructing spatial

and in variant detectors for perceptual patters. This resembles the major difficult part of the problem, but practical applicability is essential to prove the theoretical foundation on the noisy real world.

Neural networks have been used to perform complex functions in various fields to applications, including automation, robotics, defense, remote monitoring, and others. In recognition, most of the use of the learning vector quantization are applicable for clustering in recognition problems. In general, the emergence of neural network algorithms in the remote sensing field is of novel applicability. One of

¹ Department of Computer and Communication Systems Engineering, Faculty of Engineering, University Putra Malaysia 43400, UPM, Serdang Selangor, Malaysia, E-mail: hanan_adlan@yahoo.com, hanan@pd.jarring.my

² Head of Intelligent Systems and Robotics Laboratory (ISRL), Institute of Advanced Technology (ITMA), University Putra Malaysia, 43400 UPM, Serdang, Malaysia, E-mail: arr@putra.upm.edu

³ Department of Mathematics, Faculty of Science, University Pendidikan Sultan Idris

* Corresponding author

the best articles that reviewed applications in this area before the year 1997 is the one by Kanellopoulos (Kanellopoulos, 1997). Clustering of remotely sensed images for multi spectral data is the issue of many authors (Gualtieri, 1998; Kamgar-Parsi *et al.*, 1990; Hara *et al.*, 1994). The LVQ is successful in classifying SAR imagery, compared with the likelihood classifier (Knok *et al.*, 1991). The LVQ is developed to deal with compression of SAR imagery as well (Luttrell, 1988). Hybrid modeling involves an LVQ and multi-layer perceptron architecture for classification of remotely sensed images (Kanellopoulos *et al.*, 1994).

From the evidence above, it is clear that most of the work has concentrated on classification problems. In this paper, performance of a learning vector quantization network has been investigated. The issues focused on the use of invariants and spatial instances, the effect of the number of hidden units, and the learning rate parameter on the recognition ability of the work architecture for satellite imagery. Spatial and invariants detectors for patterns form the original instances for the recognition task.

The rest of the paper is organized as follow; section 2 describe the LVQ, section 3 give the theoretical foundation, while section 4 develops the architecture, section 5 discusses the results, and section 6 concludes the investigation.

Linear Vector Quantization (LVQ)

A learning vector Quantization (LVQ) is a method for training competitive layers in a supervised manner. The model employs a self-organizing network approach that uses training vectors to recursively tune placement of the competitive hidden unit that represent categories of the inputs. Once the network is trained, an input vector is categorized as belonging to the class represented by the nearest hidden unit. The hidden units may be thought of as having inhibitory connections between each other so that the unit with the largest input wins and inhibits all the other units to such an extent that only the winning unit generates an output.

Each hidden unit represent a point in an N-dimensional space. The outputs of the hidden units are based on the proximity of the input vector, and the final output is determined by the weights of the linear output layer. The linear layer transforms the competitive layer's classes into target classifications defined by the user

Theoretical Foundation

Instances for the recognition process originate from moment's in variants, block pixel variation and the standardized variance, some of the co-occurrence matrices given in this section, which also gives an introductory explanation to the Invariant Recognition.

Invariant Recognition

Invariant features are used to cope with rotation, scale, and orientation variations. This section highlights a simple idea about the invariance problem. For more depth on the subject, one can refer to Wood (1996). Assume a set V of possible pattern to be a function of set X . For each pattern $f \in V$, a classification $c(f)$ exists. The pattern recognition problem is to construct a system which takes as input an element f of V and gives an output $s(f)$ such that $s(f) = c(f)$ for all $f \in V$. The classification problem assumes a group G act on set X , which implies that G acts on V . For any $g \in G$, $f \in V$ by

$$(gf)(x) = f(g^{-1}x) \quad \forall x \in X \quad (1)$$

Given a geometrical group G , closed under translations, rotations, scaling, etc., the classification of a pattern is in variant under the group

$$c(gf) = c(f) \quad \forall g \in G, f \in V \quad (2)$$

Learning can be done either by assuming an invariance group G to be known a priori (Wood, 1996); the other assumes unknown invariants Webber (1996).

The problem in pattern recognition aims to build constraints on the function s corresponding to those on c chat is

$$s(gf) = s(f) \quad \forall g \in G, f \in V \quad (3)$$

The invariant approach should be able to improve the generalization of any given

classifier. Two approaches to deal with the invariance problems are highlighted: the classification process can be broken down into distinct phases of invariant feature extraction and feature classification. A set of invariant features has to be found. Then the patterns are classified from the information given in the feature vectors, symmetries of the actual pattern set need not be considered. Any conventional pattern recognition technique can be applied for classification. Another approach is to combine these two phases by constructing an invariant, which is parameterized. The parameters of the invariant can be adapted in order to perform the desired classification.

Momen's Invariants

Momen's invariants are obtained by taking quotients and powers of moments. A moment is a weighted sum of the pattern $f(x)$ over the input field, with weights equal to some polynomial in x .

For non-negative integers p, q the $(p + q)$ th order moments of a pattern $f(x, y)$ defined on the plane R^2 are given by

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (4)$$

$f(x, y)$ is assumed to be a piecewise continuous function which is non-zero in only a finite part of the plane, p and q are non-negative integers. The discrete form is given by

$$m_{pq} = \sum_{i=1}^n \sum_{k=1}^n x_i^p y_k^q f(x_i, y_k) \quad (5)$$

where f is a function taking non-zero values in the range $\{x_1, \dots, x_n\} \times \{y_1, \dots, y_n\}$.

The centroid (x_0, y_0) of f is defined by

$$x_0 = (m_{1,0} / m_{0,0}), y_0 = (m_{0,1} / m_{0,0}) \quad (6)$$

The central moments are defined by

$$v_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_0)^p (y - y_0)^q f(x, y) dx dy \quad (7)$$

Moment of the centralized image are translation invariant.

The normalized central pq th moments are

$$\mu_{pq} = \frac{v_{pq}}{(v_{00})^r}$$

where $r = 1 + (p + q) / 2$, and $p + q \geq 2$ (8)

The normalized moments are also scale invariants. From these functions, Hu (1962) derived seven moment's invariants, which are also translation, scale, and rotation invariants; the first five moments used in the systems are given below:

$$\varphi_1 = \mu_{20} + \mu_{02} \quad (9)$$

$$\varphi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \quad (10)$$

$$\varphi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad (11)$$

$$\varphi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \quad (12)$$

$$\varphi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(3\mu_{21} - \mu_{03}) [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

The Block Pixel Variation and the Standardized Block Variance

In image recognition application, characteristics of a certain pattern may well resemble the characteristic of others. Examples are the mean, mean square error, and the histogram (Looney, 1997). The spatial rate of frequencies of the gray level signal in both horizontal and vertical directions. Features should possess as much information to distinguish between classes; moreover, they should be invariant to scale, translation, rotation, and reflection. The block pixel variation and the standardized block variance give a robust measure for the variations within blocks of the image.

The block pixel variation computes the 8-neighbors directional derivatives $d_{ij}(r)$ with respect to the r th neighbor in the gray scale; the pixel variation is given by

$$v_{ij}^2 = \frac{1}{8} \sum_{(r=1, 8)} d_{ij}^2(r) \quad (14)$$

The block pixel variation is then given by

$$v_b^2 = \frac{1}{\mu_b^2} \sum_j v_{ij}^2 \quad (15)$$

where μ_b is the block mean pixel variation.

The standardized block variance is given by

$$\sigma_b^2 = \sum_{ij} (P_{ij} - \mu)^2 / \mu^2 \quad (16)$$

Where $\mu = \frac{1}{M \times N} \sum_{ij} P_{ij}$ for an image of size $M \times N$, and P_{ij} are the pixel values ($1 \leq i \leq N$), ($1 \leq j \leq M$).

Co-Occurrence Matrices

The co-occurrence method describes the second-order image statistics, and works well for a large variety of textures. Good properties of the co-occurrence method are the description of spatial relations between pixels, and invariance to monotonic gray-level transformation (Milan *et al.*, 1999). The method based on the repeated occurrence of some gray-level configuration in the texture, this configuration varies rapidly with distance. The symmetric matrix $\mathbf{p}_{\phi,d}(a,b)$ gives the configuration of the relative frequencies, where d and ϕ represent distance and direction, respectively. The computation processes of the co-occurrence matrices consider only direct neighbors, and treat the relation as symmetric (without orientation).

The following co-occurrences are computed for each block:

Energy, (angular second moment) defined as

$$\sum_{a,b} \mathbf{P}_{\phi,d}^2(a,b) \quad (17)$$

Entropy, given by

$$\sum_{a,b} \mathbf{P}_{\phi,d}^2(a,b) \log_2 \mathbf{P}_{\phi,d}(a,b) \quad (18)$$

Contrast, defined by

$$\sum_{a,b} |a,b|^k \mathbf{P}_{\phi,d}^k(a,b) \quad (19)$$

Correlation, defined as

$$\frac{\sum_{a,b} [(a,b) \mathbf{P}_{\phi,d}^2(a,b)] - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (20)$$

Where μ_x, μ_y are means, and σ_x, σ_y are standard deviations.

The Architecture

The architecture used in the experiments consists of 24 input neurons, varying hidden units and three output neurons. Weights were updated after

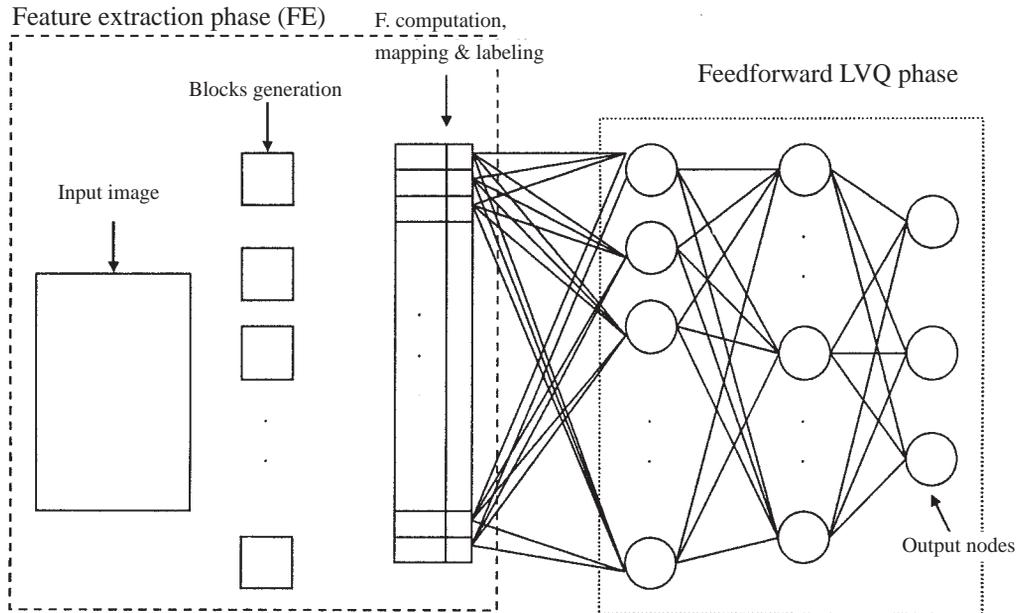


Figure 1. Feed forward LVQ neural network architecture.

the presentation of each training vector using a variable learning rate at momentum to speed training. The learning rate was reduced dynamically after each epoch and processing continued using random starting initialization.

The architecture, which is composed of two phases, is depicted in Figure 1. The two phases are a feature extraction phase (FE), and a feedforward neural network phase. The FE accepts a two dimensional input image, generate a $n \times n$ sub blocks of the image, performs feature computation, mapping and labeling for each block. Output of the FE phase are the input to the second phase. Neurons in the input layer are activated by the FE detectors as inputs, during training the network takes the inputs and their corresponding labels. Placement of the features is performed randomly. For testing, the network accepts the feature detectors as input, and produces the class identifier.

Results and Discussion

A learning vector quantization has been examined for the recognition of three categories presented by water, buildings and network, and vegetations. 24-feature detector representations of satellite image pattern were extracted from the 3×3 block composed from the original scene. All coefficients were mapped to the bipolar interval of [-1, 1] using equation (21). For each of the three patterns, the training set and the evaluation set were used for training and testing. This is illustrated in Figure 2.

The test is performed on the ability of the neural architectures to classify an input vector as one of the three categories on which it was trained. Performance of the network was measured in term of the recognition rates for both the training and the test categories.

Two major factors affect the behavior of

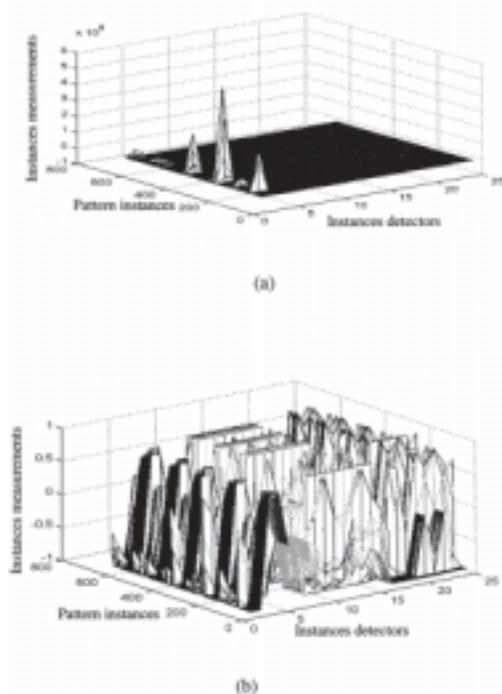


Figure 2. Visualization of the effect of mapping the pattern instances (a) original instance; (b) tranformed instances.

Table 1. Effect of different hidden units (evaluation set).

Neural network	Number of hidden units	Recognition rate* (%)
LVQ	200	87.81
	150	91.44
	100	84.88
	50	84.54
	10	76.05

* Average of 6 networks

Table 2. Effect of different learning rates (evaluation set).

Neural network	Learning rates*	Recognition rate* (%)
LVQ	200	87.81
	150	91.44
	100	84.88
	50	84.54
	10	76.05

* Average of 6 networks

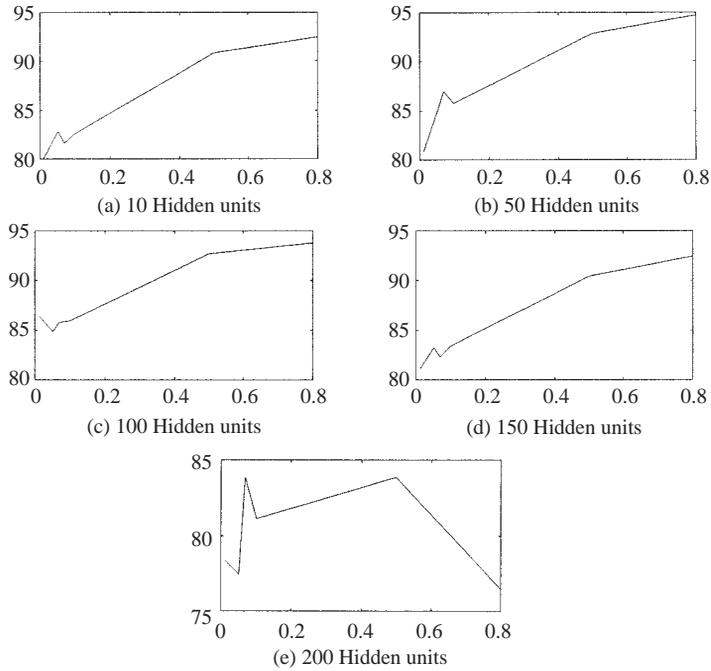


Figure 3. Effect of different hidden units on the recognition ability of LVQ network (training set), y-axis: recognition rates, x-axis: learning rates.

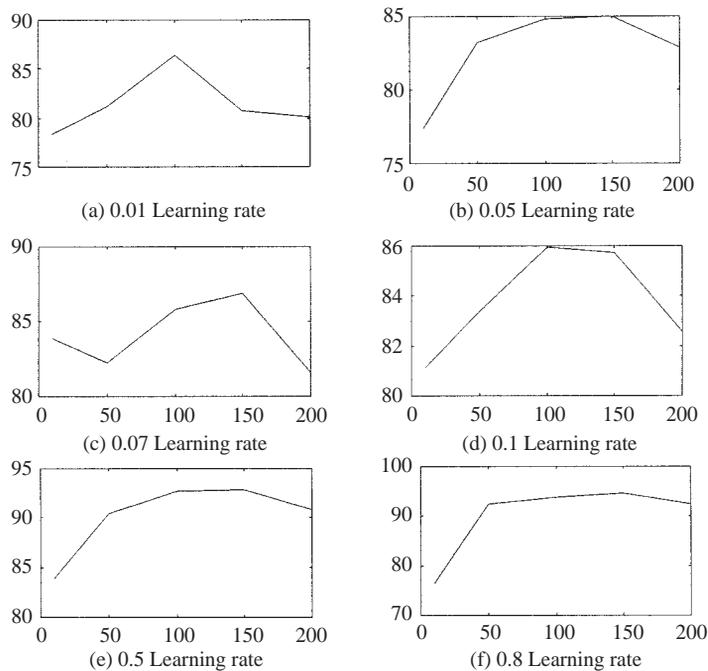


Figure 4. Effect of different learning parameter on the recognition ability of LVQ network (training set), y-axis: recognition rates, x-axis: hidden neurons.

the LVQ network, the number of hidden units, and the learning rate. Figure 3 and 4 give the effect of the number of hidden units, and the learning rate parameter, respectively for the training category. For the evaluation category Table 1 and 2 summarize the average of the networks for both the effect of the hidden units and the learning rate parameter.

Higher recognition rates of the training set Figure 3 (a-e) are achieved with the increase in the number of hidden units as seen in Figure 3 (a-c), and then stated to decrease Figure 3 (d-e). However, generalization degraded with large number of hidden units and the amount of computation required increased as well. The effect of the hidden units for the evaluation set Table 1 is given as average of the 6 networks. Again, the recognition rates are increasing as the number of hidden units increase, and started to decrease for higher units of about 200 as shown in Table 1. The network generalization seemed to decrease as the hidden units were added.

Figure 4 gives a visualization of different learning rates (a-f); there is a systematic trend in the graphs. The best recognition rates are achieved with 100, and 150 hidden neurons for different learning rates. The evaluation of the network results in Table 2 is displayed as an average of 5 networks, where the best result is obtained for the higher learning rates.

Conclusion

This paper demonstrates the use of invariants detectors coupled with spatial ones to represent the pattern for an LVQ network. A simple architecture that consists of two layers produces the significant recognition rates. Varying hidden units and learning rates have been examined for investigating the effect of such parameters on the recognition of certain pattern presented in the image. The LVQ network requires a short training time to perform the network computation. Other neural nets have to be examined in order to obtain a better recognition, which is the issue for further research.

References

Gualtieri, J.A. (1988). Goddard researchers

simulate neural networks using parallel processing. NASA Information Systems Newsletter, 15:12-15.

- Hara, Y., Atkins, R.G., Yueh, S.H., Shin, R.T., and Kong, J.A. (1994). Application of neural networks to radar image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 32(1):100-109.
- Kamgar-Parsi, B., Gualtieri, J.A., and Devaney, J.E. (1990). Clustering with neural networks. *Biological Cybernetics*, 63:201-208.
- Kanellopoulos, I. (1997). Use of neural networks for improving satellite image processing techniques for land cover/land use classification. Italy: European Commission, Joint Research Center. Available from: <http://europa.eu.int>. Accessed June, 2001.
- Kanellopoulos, I., Wilkinson, G.G., and Chiuderì, A. (1994). Land cover mapping using combined Landsat TM imagery and textural features from ERS-1 Synthetic Aperture Radar Imagery. In: *Image and Signal Processing for Remote Sensing*. Desachy, J., ed., p. 332-341.
- Knok, R., Hara, Y., Atkins, R.G., Yueh, S.H., Shin, R.T., and Kong, J.A. (1991). Application of neural networks to sea ice classification using polarimetric SAR image. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'91)*; June 1991; Espoo, Finland. IEEE Press, 1:85-88.
- Looney, C.G. (1997). *Pattern recognition using neural networks. Theory and Algorithms for Engineer and Scientists*' Oxford.
- Luttrell, S.P. (1998) Image compression using a neural network. *Proceedings of the International Geoscience and Remote Sensing Symposium, (IGASS'88)*; September, 1998; Edinburgh, Scotland. ESA Publications Division, p. 1,231-1,238.
- Milan, S., Hlavac, V., and Boyle, R. (1999). *Image processing, analysis, and machine vision*. ITP.
- Webber, C. (1994). Self-organization of transformation invariant detectors for

constituents of perceptual pattern. Wood, J. (1996) Invariant pattern recognition:
Network: Computation in Neural Syst., A review' pattern recognition. Elsevier
5:471-496. Science, 29(1):1-17.