

An Heuristic Pattern Correction Scheme for GRNNs and Its Application to Speech Recognition

Tetsuya Hoya and Anthony G. Constantinides
Signal Processing and Digital Systems Section,
Dept. of Electrical and Electronic Engineering,
Imperial College of Science, Technology, and Medicine,
University of London, SW7 2BT, U.K.
email: t.hoya@ic.ac.uk

Abstract

In an on-line learning environment where optimal recognition performance over the newly encountered patterns is required, a robust incremental learning procedure is necessary to re-configure the entire neural network without affecting the stored information. In this paper, an heuristic pattern correction scheme based upon an hierarchical data partitioning principle is proposed for digit word recognition. This scheme is based upon General Regression Neural Networks (GRNNs) with initial centroid vectors obtained by graph theoretic data-pruning methods. Simulation results show that the proposed scheme can perfectly correct the mis-classified patterns and hence improves the generalisation performance without affecting the old information. Moreover, it is also established that the initial setting of Radial Basis Functions (RBFs) based upon graph theoretic data-pruning methods yields better performance than those obtained by k -means and Learning Vector Quantisation (LVQ) methods.

1 Introduction

Incremental learning is an efficient learning mechanism for neural networks that adds new information during training without re-initialisation of the entire network. The development of promising incremental learning methods is therefore an issue with great interest in the study of neural networks. Probabilistic Neural Networks (PNNs) [1] and

Generalised Regression Neural Networks (GRNNs) [2] share a special property, namely they do not require iterative training since the weight vector between the Radial Basis Functions (RBFs) and the output unit can be fixed as the target vector. This attractive property is particularly useful in on-line supervised learning [3], as incremental training may be achieved without affecting the stored information [4, 5].

In this paper, we propose an heuristic pattern correction scheme for GRNNs and apply it to the correction of the mis-classified patterns for digit word recognition. The proposed method, unlike the Parzen classifier based methods in [6, 7], takes an *instance – based* approach with the aid of an hierarchical data partitioning mechanism, which eliminates the need for statistical density approximation and its associated considerable mathematical complexity. Moreover, we also compare the correction performance with five different initial subset settings based respectively upon the k -means, LVQ, Vertex-Chain [8], List-Splitting [8], and SST-Splitting [8] methods. In the simulation study, complete correction was achieved with a relatively small number of RBFs, and it was observed that the subset settings chosen by the graph theoretic data-pruning algorithms yield slightly better recognition performance than those with the k -means and Learning Vector Quantisation (LVQ) methods.

2 The Pattern Correction Scheme

In applications of neural networks to pattern classification tasks, a complete re-initialisation procedure is normally required to learn the newly added patterns or to retrain the mis-classified patterns. In an on-line learning environment, this is time-consuming especially when the size of the original training data set is very big and an instant pattern correction scheme on the in-coming data is therefore necessary. To meet this requirement, the proposed pattern correction scheme takes an hierarchical data partitioning approach and employs the special property of GRNNs, namely no iterative learning procedure is required for the newly added RBFs.

The pattern correction is performed in an iterative fashion and is performed according to the following:

- Step 1. Initialise the iteration count for the correction, $cnt = 1$.
- Step 2. Test the performance of the original GRNN with all the testing patterns.

- Step 3. Collect the mis-classified patterns. Then choose heuristically the patterns for new RBFs among the testing patterns, and add them into the GRNN (i.e., Network Growing). The weight vector between the new RBF and the output neurons is fixed identical to the target vector of the corresponding mis-classified pattern.
- Step 4. Test the performance of the grown GRNN with the entire testing pattern set.
- Step 5. If there is no mis-classification, then terminate. Otherwise $cnt \leftarrow cnt + 1$, then return to Step 2.

In Step 3 above, an heuristic method is used to choose the additional patterns for the new RBFs. For the English digit word recognition task, the number of categories N_c is already known ($N_c = 10$) and therefore in this paper the maximum number of RBFs to be added to the network in an iteration is fixed identical to that number of categories, in order to keep the growing network always in a “well-balanced” shape in terms of its generalisation capacity. In Step 3 above, the radii values of RBFs should also be updated since the pattern space represented by the centroid vectors would be changed. The radii setting of RBFs is described later in this paper.

3 Network Setting for Pattern Recognition

A fully connected Multi-Layered GRNN (ML-GRNN) is used, which has 256 input neurons, M RBFs, and 10 output neurons. The number of RBFs grows during the iterative correction process. The target vector for pattern i is given as a vector of indicator functions:

$$\mathbf{T}_i = (\delta_1, \delta_2, \dots, \delta_{10}),$$

$$\delta_j = \begin{cases} 1 & \text{if pattern } i \text{ belongs to the category (digit)} \\ j - 1 & (j = 1, 2, \dots, 10) \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

With the setting above, the topology of the ML-GRNN with 10 output units can be seen as a set of 10 sub-nets with a decision unit as illustrated in Fig. 1, since the weight having the value 0 can be removed from the network. On the right hand side of the figure, each sub-net is viewed as a collection of RBFs which represents the entire pattern

space for a single category. With the network on the right, the final decision is therefore made following the “winner-takes-all” strategy.

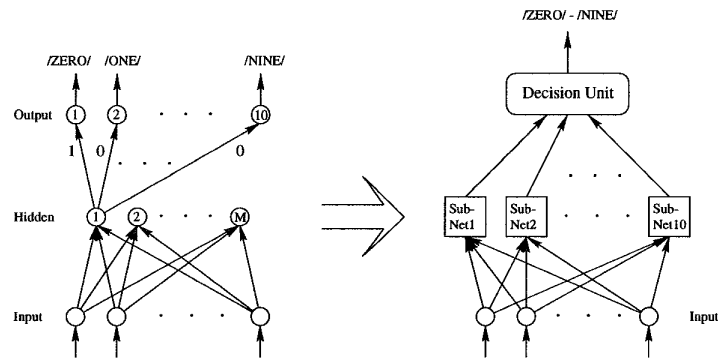


Figure 1: Illustration of Topological Equivalence Between the ML-GRNN With M Hidden and 10 Output Units and the Assembly of the 10 Distinct Sub-Nets

3.1 Radii Setting

The setting of radii values is also a significant factor for the design of RBF-NNs and such determination is still one of the open issues [9, 10]. In the simulation work, we have investigated the individual setting of radii values using 1-nearest neighbour [11], however, the recognition performance using this technique did not yield better results than the radii setting with fixed values. In this paper, fixed radii values for the respective RBFs are therefore used and set equal according to the following modified radii setting found in [10]:

$$\sigma = \frac{d}{N\sqrt{2M}}, \quad (2)$$

where d is the maximum Euclidean distance between the centroid vectors, M is the number of RBFs, and N is the number of units in the output layer of the ML-GRNN. In this paper, the radii values are updated during the network growing phase according to Eqn. 2.

4 Simulation Study

4.1 Pattern Set

In the experiment, the data set used for English digit word recognition is a volume of the SFS database [12], containing 500 utterances of the digit from /ZERO/ to /NINE/ recorded by two male and three female speakers. The volume is then arbitrarily partitioned into three distinct sets: the set for training, testing, and the set used for the performance test. The training set consists of a total of 250 patterns which are gathered evenly from five different speakers of five utterances for each digit. The testing set, likewise, consists of evenly selected 100 patterns. The rest (150 patterns) are thus used for the performance test (i.e., the unknown data). Each utterance is sampled at 20kHz, and is converted into a feature vector with a normalised set of 256 data points which are obtained by the LPC-Mel-Cepstral analysis. The feature vector is therefore used as the input vector of the GRNN.

4.2 Initial Choice of RBFs

The initial choice of centres was performed by k -means, LVQ, and the three graph theoretic pruning methods proposed in [8], i.e., Vertex-Chain, List-Splitting, and SST-Splitting methods. The original GRNN was composed with 20-RBFs, which yielded only a poor recognition performance as in Table 1.

Method	Total Mis-Classifications
k -means	60/150 (60.0%)
LVQ	76/150 (49.3%)
Vertex-Chain	35/150 (76.7%)
List-Splitting	38/150 (74.7%)
SST-Splitting	33/150 (78.0%)

Table 1: A Comparison of Recognition Performance Over the Unknown Data With the Original GRNN (initially set up with 20-RBFs)

4.3 A Further Optimisation for the Correction Scheme

From above, it is already known that utterances are collected from five different speakers for each category (digit) to compose the speech

data set, it is possible to make the choice become less arbitrary. In this paper, by using the *a priori* knowledge, a further modification in the following is therefore considered:

For each category (digit), do the following steps:

- Step 1. If there is no mis-classification, skip this category and move on to the next category. Otherwise go to the next step.
- Step 2. Enumerate the number of mis-classifications per person for this category. Then sort the list and find the person whose mis-classification rate is maximum.
- Step 3. Take arbitrarily one pattern among the mis-classified patterns of the person found in the previous step, and add it to the network as a new RBF.

4.4 Experimental Results

The experiment was performed with the parameter setting and the optimised pattern correction scheme described above. Fig. 2 and 3 show respectively the transition in the total number of RBFs in the network and that of mis-classified patterns out of the testing data set achieved by the GRNN during the iterative pattern correction process. In Table 2, a comparison of recognition performances over the unknown data with the five different methods after completing the mis-classification correction is shown. In comparing Table 2 with Table 1, overall im-

Method	Total Mis-Classifications	Num. of RBFs after Correction
<i>k</i> -means	19/150 (87.3%)	43
LVQ	31/150 (79.3%)	49
Vertex-Chain	15/150 (90.0%)	41
List-Splitting	18/150 (88.0%)	40
SST-Splitting	14/150 (93.3%)	47

Table 2: A Comparison of Recognition Performance Over the Unknown Data With the GRNN After the Pattern Correction

provement in recognition performance with the proposed pattern correction method is obtained for all the pruning algorithms. In Table 2,

it is also observed that some of the graph theoretic methods perform better than k -means or LVQ method.

5 Conclusion

In this paper, we have proposed an heuristic pattern correction method specialised for digit word recognition based upon GRNNs. The proposed network generalisation method enables us to correct perfectly the mis-classified patterns with relatively small numbers of RBFs and is considered to be suitable for application to a strict security service where recognition performance without failure over a specific pattern set is required. In the experiment, the data partitioning point between the training and the testing patterns, however, was arbitrarily determined, thus whether the network grown by the presented correction scheme yields the best performance over the unknown data set is still not guaranteed. Related to this problem, cross-validation [9] using graph theoretic partitioning technique is also under investigation.

References

- [1] D. F. Specht, "A General Regression Neural Network", IEEE Trans. on Neural Networks, Vol. 2(6), pp.568-576, 1991.
- [2] P. D. Wasserman, "Advanced Methods in Neural Computing – in Chapter 8, Radial Basis-Function Networks", pp.147-176, Van Nostrand Reinhold, New York, 1993.
- [3] C. P. Lim and R. F. Harrison, "An Incremental Adaptive Network for On-line Supervised Learning and Probability Estimation", Neural Networks, Vol. 10, No. 5, pp.925-939, 1997.
- [4] L. Fu, "Incremental Knowledge Acquisition in Supervised Learning Networks", IEEE Trans. on SMC. Part A, Systems and Humans, Vol. 26, No. 6, pp. 801-809, Nov. 1996.
- [5] M. J. L. Orr, "Introduction to Radial Basis Function Networks", Material on Internet found at www.cns.ed.ac.uk, Centre for Cognitive Science, University of Edinburgh, Apr. 1996.
- [6] P. A. Devijver and J. Kittler, "Pattern Recognition: A Statistical Approach", Prentice Hall International, 1982.

- [7] M. A. Kraaijveld and R. P. W. Duin, "Generalization Capabilities of Minimal Kernel-Based Networks", Proc. of Int. Joint Conf. on Neural Networks , Seattle, 1991.
- [8] T. Hoya, "Graph Theoretic Techniques for Pruning Data and Their Applications", to appear in IEEE Trans. on Signal Processing.
- [9] C. M. Bishop, "Neural Networks for Pattern Recognition", Oxford Univ. Press, 1996.
- [10] S. Haykin, "Neural Networks: A Comprehensive Foundation", Macmillan College Publishing Co. Inc., 1994.
- [11] A. Saha and J. D. Keeler, "Algorithms for better representation and faster learning in radial basis function networks", In: "Advances in Neural Information Processing Systems 2", Ed. D. S. Touretzky, pp.482-489, San Mateo, CA: Morgan Kaufmann, 1990.
- [12] M. Huckvale, "Speech Filing System Vs3.0 – Computer Tools For Speech Research", University College London, Mar. 1996.

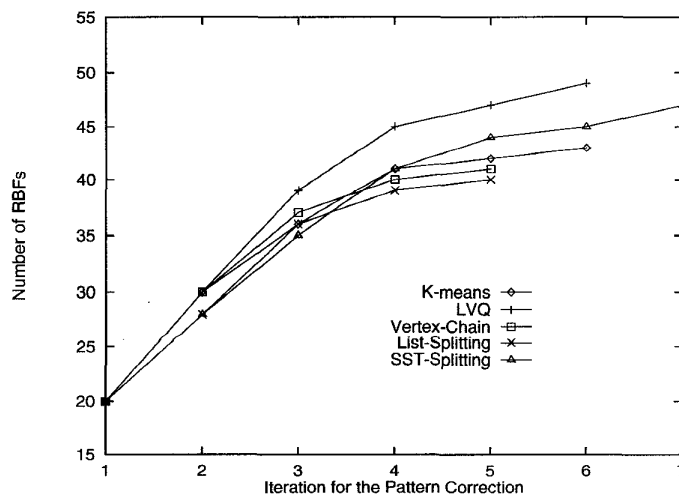


Figure 2: Transition of the Number of RBFs during the Pattern Correction

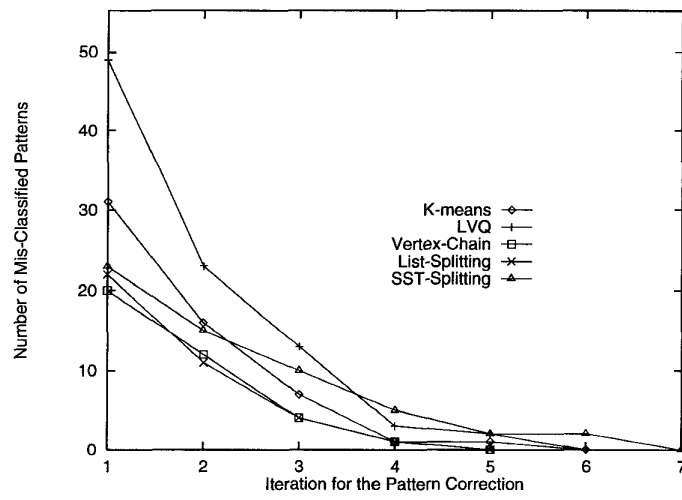


Figure 3: Total Number of Mis-Classified Patterns Out of the Testing Data Set during the Pattern Correction