行政院國家科學委員會專題研究計畫　成果報告

類神經網路與樹狀自動辨認機於震測圖型識別之研究

計畫類別：個別型計畫
計畫編號：NSC91-2213-E-009-095-
執行期間：91年08月01日至92年07月31日
執行單位：國立交通大學資訊科學學系

計畫主持人：黃國源

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國　92年10月24日

行政院國家科學委員會專題研究計畫成果報告

類神經網路與樹狀自動辨認機於震測圖型識別之研究

The Study of Neural Networks and Tree Automata for Seismic Pattern Recognition

計畫編號：NSC 91-2213-E-009-095
執行期限：91 年 8 月 1 日至 92 年 7 月 31 日
主持人：黃國源　交大資訊科學系　kyhuang2@cis.nctu.edu.tw

一、中文摘要

本研究提出一個樹狀結構的自動機辨識系統，用來辨識在震計記錄圖中的的震測圖型。類神經網路的多層感知器被用來識別震測圖型中的子圖型，這些子圖型可以建構出震測圖型的樹狀結構。吾人利用三種由下往上結構保留可校正錯誤樹自動辨識機來辨識這些樹狀結構的震測圖型，此外提出了新的由上往下可校正錯誤的樹自動機，用來辨識非結構保留的樹狀結構的震測圖型。我們應用於模擬的與真實的亮點震測圖型，其辨識結果有助於探油的震測解釋。

關鍵詞：樹自動機，樹狀文法，多層感知器，震測圖型。

Abstract

We propose a tree automaton system for the recognition of structural seismic patterns in a seismogram. Multilayer perceptron neural network is used for the identification of subpatterns, with which a tree representation of the structural seismic pattern is constructed. We use three kinds of modified bottom-up structure preserved error correcting tree automata (SPECTA) to recognize the tree representation of pattern, and propose a new top-down error correcting tree automaton (ECTA) to recognize non-structural preserved pattern. In the experiments, the system is applied to the simulated and real seismic bright spot patterns. The recognition result can improve seismic interpretation.

Keywords: tree automata, tree grammar, multilayer perceptron, seismic patterns.

二、緣由與目的

The methods of syntactic pattern recognition have been adopted for the recognition of seismic patterns (Huang et. al., 1987, 1992). Most of the papers focused on one dimensional pattern string distance computation, finite-state grammar and automaton, attributed context-free grammar and automaton for the recognition of one dimensional (1-D) seismic wavelets and reflector horizons in the seismogram. But the 1-D pattern grammar is not easy to describe the 2-D or 3-D seismic pattern.

If we use tree grammars and automata for the description and recognition of seismic patterns, then the most critical problem in tree automaton is in the tree construction of a pattern. In real seismic data of bright spot pattern, there is a lot of interference. There may not be connected between reflection horizon and horizon. It becomes difficult to construct a tree representation for a complete pattern. Then tree automaton becomes infeasible.

If we use neural network method. The famous Fukushima's model recognizes the segments first, then subpatterns, larger subpatterns, …, finally the whole pattern. It is a neural computing method of hierarchical (structural) recognition. Because there are many layers in the neural network, the learning and the computation are very complex.

Here we adopt the advantages of tree automaton and Fukushima's model. Instead of recognizing the horizon segments, we use neural network in the recognition of subpatterns of seismic bright spot pattern. Then from the relation between subpattern and subpattern, we construct a tree for a seismic bright spot pattern. Finally the tree automaton can parse the tree. Because multilayer perceptron neural network has good gradient descent training algorithm, we use it for the recognition of subpatterns.

So we use tree grammars and automata for the description and recognition of the bright spot seismic pattern. The proposed system contains two parts: the training and the recognition parts, as shown in Figure 1. In the training part, the desired seismic pattern is transformed into the corresponding tree representation, from which the tree grammar (Fu, 1982) and the automaton can be inferred. The recognition of patterns follows three steps: preprocessing, pattern representation, and error-correcting tree automata.

A bright spot structure after preprocessing is shown in Figure 2. Using the eight directional Freeman's chain codes in Figure 3 as primitives, the tree representation of the bright spot structure can be constructed as that shown in Figure 4, and can infer tree grammar as follows.

Tree grammar: G_t = (V, r, P, S), where

V=set of terminal and nonterminal symbols

={$, 0, 5, 7, @, S, A, B, C, D, E, F, G, H, I, J},

V_T = the set of terminal symbols = { $, 0, 5, 7, @ },

$: the starting point (root) of the tree,

@: the neighboring segment has already been expanded

(assuming a top-down, left-right expansion),

S: the starting nonterminal symbol,

r: r(5)=r(7)={2,0}, r($)=2, r(@)=0, r(0)=0,

and P:

```
(1)S→$    (2)A→5    (3)B→7    (4)C→5     (5)D→0  (6)E→@
    / \       / \       / \       / \
   A   B     C   D     E F       G  H
(7)F→7    (8)G→5    (9)H→0    (10)I→@    (11)J→7
    / \
   I   J
```

The tree production rules P can derive trees. Each tree is corresponding to its seismic pattern.

In the recognition of subpatterns, we break the seismic bright spot pattern into five subpatterns as shown in Figure 5. Seven moments are extracted from each subpattern, which has been shown to be invariant to translation, rotation, and scale changes (Hu, 1962).

Seven moments from each subpattern are input to the multilayer perceptron neural network (Rumelhart et al., 1986) in Figure 6 for subpattern classification. Use the relative positions of subpatterns, the tree representation of the seismic pattern can be constructed. The tree is then parsed by the error-correcting tree automaton into the correct class.

Due to noise, distortion, and interference of the wavelets at the junctions, the seismic pattern is broken into many horizon segments, the tree structure may be deformed with substitution, deletion, or insertion errors. If the tree structure is preserved with only substitution error, bottom-up structure preserved error-correcting tree automata (SPECTA) can be applied in the recognition of the pattern. Modified weighted minimum-distance SPECTA, modified maximum-likelihood SPECTA, and fuzzy SPECTA are used in the recognition of simulated seismic bright spot pattern.

However the deformation may lead to non-structural preserved tree structure. We propose a top-down minimum-distance ECTA which can recognize tree structure with substitution and deletion errors in the real seismic bright spot pattern.

三、結果與討論

In the simulation experiment, three SPECTA are used in the recognition of simulated bright spot pattern. The recognition results are good.

In the real data experiment, the real seismic data at Mississippi Canyon is shown in Figure 7(a). After preprocessing, the result is shown in Figure 7(b). Preprocessing includes thresholding, compression, and thinning. After subpattern recognition,

Figure 7(c) shows the center positions of subpatterns. From the relation of the center positions of subpatterns, the tree representation of the extracted bright spot pattern is shown in Figure 7(d). The tree in Figure 7(d) is parsed by the proposed top-down ECTA. The result in Figure 8 shows that the error distance is 5. If the threshold is set to 5 or smaller, then the tree is accepted as bright spot. However the tree in Figure 7(d) cannot be parsed by a bottom-up SPECTA, because the tree structure is not preserved.

The tree representation is quite critical in the tree automaton. Due to noise, distortion, and interference of the wavelets at the junctions, the seismic pattern is broken into many horizon segments. Without complete tree, the seismic pattern can not be parsed and recognized. We have proposed a system to combine subpattern recognition with neural network and whole pattern recognition with tree automata such that syntactic approach can work on seismic bright spot recognition. The recognition results can improve seismic interpretation.

四、成果自評

研究內容與原計畫相符程度: 100%
達成預期目標情況: 100%
研究成果的學術或應用價值: 建立樹狀結構自動機系統於震測圖型之辨認
是否適合在學術期刊發表: 是
主要發現或其他有關價值: 可用於中文字的強健性之辨認

五、參考文獻

Barrero, A., 1991, Inference of tree grammars using negative samples: Pattern Recognition, 24, 1-8.

Fu, K. S., 1982, *Syntactic Pattern Recognition and Applications*: Prentice-Hall.

K. Fukushima, 1989, “Analysis of process of visual pattern recognition by the neocognitron,” Neural Network, Vol. 2, pp. 413-420.

Hu, M. K., 1962, Visual pattern recognition by moment invariants: IRE Trans. Info. Theory, IT-8, 179-187.

Huang, K. Y., Fu, K. S., Cheng, S. W., and Lin, Z. S., 1987, Syntactic pattern recognition and Hough transformation for reconstruction of seismic patterns: Geophysics, 52, 1612-1620.

Huang, K. Y., and Leu, D. R., 1992, Modified Earley parsing and MPM method for attributed grammar and seismic pattern recognition: Journal of Information Science and Engineering, 8, 541-565.

Kou-Yuan Huang, 2001, “Neural network for robust recognition of seismic patterns,” presented at IJCNN, Washington, D.C., July 14-19.

Kou-Yuan Huang, 2003, *Neural Networks and Pattern Recognition*: Weikeg Publishing Co., Taipei, Taiwan.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986, Learning internal representations by error propagation: in Rumelhart, D. E., and McClelland, J. L., Eds., PDP, 1, M.I.T. Press, 318-362.

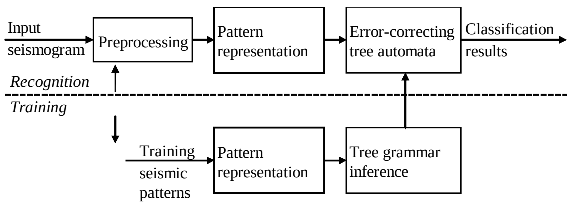


Fig. 1. Tree automaton system for seismic pattern recognition.

Fig.2. Bright spot pattern after preprocessing.

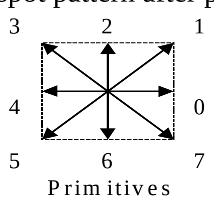


Fig. 3. Eight directional Freeman's chain codes.

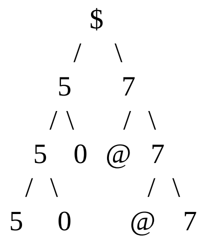


Fig. 4. Tree representation of bright spot pattern.

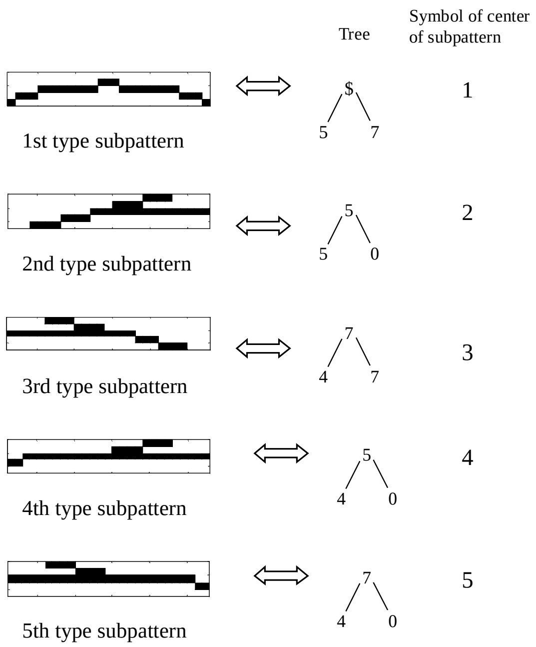


Fig. 5. Five subpatterns of bright spot and their tree representations.

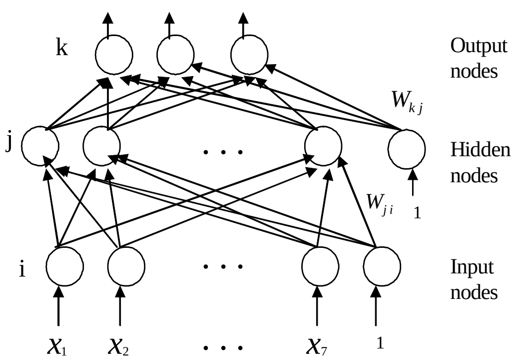


Fig. 6. Multilayer perceptron.

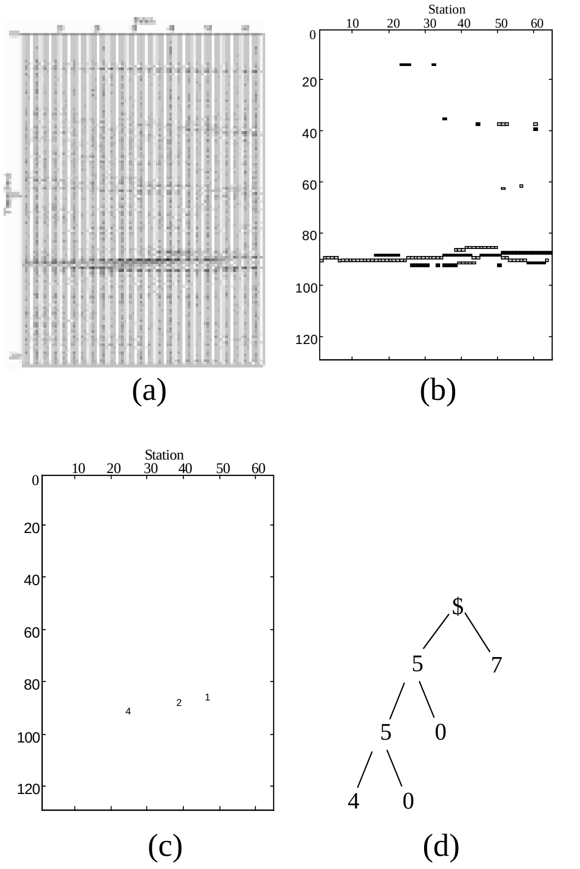
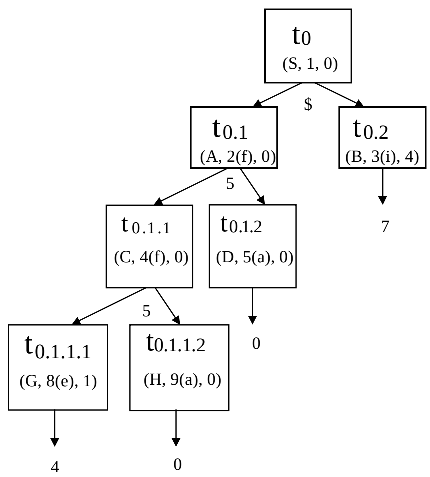


Fig. 7. (a) Seismogram at Mississippi Canyon,
(b) Peak seismogram after preprocessing,
(c) Center positions of subpatterns.
(d) Tree representation of pattern.



Minimum-distance c = 0+0+4+0+0+1+0 = 5

Fig. 8. Parsing of tree in Figure 7(d) using top-down ECTA.