寄 書

パターン認識を応用した自動 地震識別法

(財)鉄道総合技術研究所 防災技術研究部*

束田進也†

気象庁 地震火山部**

大竹和生

An Event Discrimination Method Applied
Pattern Recognition

Shin'ya TSUKADA

Disaster Prevention Technology Division,
Railway Technical Research Institute,
Hikari-cho 2-8-38, Kokubunji,
Tokyo 185-8540, Japan

Kazuo OHTAKE

Seismological and Volcanological Department,
Japan Meteorological Agency,
Ote-machi 1-3-4, Chiyoda,
Tokyo 100-8122, Japan

(Received July 27, 2000;

Accepted December 11, 2000)

§1. はじめに

震源決定は地震学の諸解析の中で最も基本的な課題の一つである。気象庁をはじめ各大学や科学技術庁防災科学技術研究所などは、日本およびその周辺で日々発生している地震について、主に手作業による検測とコンピュータを用いた震源決定作業を定常業務、あるいは準定常的作業として続けてきている。

近年この作業は観測網の高密度化、高品位化に伴って非常に膨大なものになってきている。例えば、気象庁では1960年代にコンピュータによる震源決定作業が導入された当時、震源決定数は年間数100個のオーダーであったものが、現在は年間100,000個近くに達しようとしている(Fig. 1)。この数は今後の観測網整備や震源決

定能力の向上、またより微小な地震の震源まで決定することに伴ってさらなる増加が予想される。現状のような手作業主体の検測・震源決定作業は、作業者の増員、あるいは自動化なしではもはや限界に達しつつある。

一方、コンピュータの飛躍的な高速化と大容量化に伴い、検測・震源決定のためのさまざまな自動処理システムが開発されてきた[横田・他(1981), 松村(1989), 卜部・束田(1992), 堀内・他(1999)ら]。これらのシステムの多くは、①単独の観測点の波形を用いたトリガ検出、②地震識別、③地震波の位相自動検出、④震源決定、というアルゴリズムを用いて作成されている。中でも③、④の部分は、人間が行う検測と同程度の精度で検測が可能で高精度な震源決定が行えるように、例えばARモデルの適用など開発が積極的に進められてきた。[例えば、白井・徳弘(1979), 森田・浜口(1984), 前田(1985), TAKANAMI and KITAGAWA(1988)]

ところが①と②、つまり信号が地震か否かを判別する手法に関しては、これまで地震を取りこぼさないことに主眼が置かれ、地動ノイズによって誤った地震識別を行う可能性を容認してきた。このような状況下では位相の誤認によって本来存在しない地震を決めたり、計算が不安定になって誤った震源を求めたりして自動処理が失敗することがある。別の言い方をすれば、自動検測や震源決定の手法はほぼ完成されているにも関わらず、前段階の不適切な処理のために自動処理そのものの信頼性が損なわれていると言ってもよい。

自動検測・自動震源決定は、人間による再チェックが無くてもつねに信用できる結果が得られなければ意味がない。つまり結果が5割、あるいは8割確かだとしても残りが不正確であれば、再度すべてのデータを人間がチェックしなければならなくなってしまう。

本論文では、パターン認識の考え方を導入することによって①と②の地震識別をより正確に行い、さらに先験的に震央などの情報を知ることによって、既存の自動検測・自動震源決定の成績が向上できることを述べる。

§2. 手 法

従来のトリガ検出・地震識別の手法はグループトリガ方式といわれるもので、あらかじめ決められた領域内で振幅増加が検出された(=トリガ検出された)観測点の数が一定数以上になった時に地震が発生したとみなす。この手法は1観測点を基本単位とし、これを多点組み合わせることによって局地的なノイズによる誤った地震識別を減らすことができる。トリガ検出、および地震識別は毎秒行う演算であるため計算機に大きな負担をかけるが、この方式は演算時間を軽減できるという長所があ

* 〒185-8540 東京都国分寺市光町2-8-38

** 〒100-8122 東京都千代田区大手町1-3-4

† 連絡責任者: 電話: 042-573-7273, FAX: 042-573-7421

E-mail: tsukada@rtri.or.jp

Fig. 1. Change of yearly number of hypocenters determined by Japan Meteorological Agency. The triangles indicate major earthquakes around Japan from 1960 to 1999. A, B and C show major change of seismic network and data processing system in JMA. The number of hypocenters rise dramatically after 1990s' seismic network changes.

る。しかし観測点相互間には処理が独立で、互いに連携なしに地震識別が行われるため、複数の観測点にはほぼ同時に現れたノイズ群の場合でも地震が発生したと誤認する場合がある。これを避けるにはトリガ検出した地震波の見かけ速度を計算し、ありえない見かけ速度が求まった観測点の組み合わせを除外すればよい。ただし、この計算では観測点間の距離をそれぞれ求めたり、見かけ速度の大小を判定したりする必要があってやや煩雑となってしまう。

ところで、われわれは地震観測に多チャンネルのペンレコードを用いることがある。例えば Fig. 2 のような多チャンネルの地震波形記録の場合、地震が観測網の中で発生したのか否か、観測網のどのあたりで発生した地震なのか、大きな地震なのか否か、あるいは単なるノイズなのか否かなどをわれわれは瞬時に判断できるようになる。これは学習によって地震波の見かけ速度の概念を画像パターンイメージとして理解できたためであり、最も基本的なパターン認識の例といえる。本研究はこのパターン認識のアイデアを用い、仮想した震源から作成

する地震波（時系列）データのモデルパターンと実際に観測された波形の到着パターンを比較することによって地震の識別を行う。モデルパターンを作成する際に見かけ速度の概念が入っているために、規則性無く出現するノイズは取り除かれる。

地震波形を 2 値化するにはいくつかの手法が考えられる。今回は最も単純に連続波形の振幅がある閾値を越えた場合に「1」、越えない時には「0」を毎秒算出しビットイメージのパターンを作成した (Fig. 3)。なお、この閾値は観測点ごとに地震波形の短時間平均値 (STA) と長時間平均値 (LTA) の比がある設定した値を越えるか否かを判断基準に設けた。一般に大きな閾値の場合、微小な振幅変化ではビットが「1」にならず大きめの地震しか捉えられなくなる。逆に小さな閾値なら、微小な振幅変化も捉えられるが地動ノイズと地震とのビットパターンの識別が困難になる。今回は各観測点ごとに試行錯誤で 2.0 前後の閾値を設定した。

識別可能な地震のマグニチュードの下限はパターンの白黒決定の閾値ばかりではなく、観測点密度にも左右さ

Fig. 2. Example of multi-channel seismic waveform (offered by Earthquake Observation Center, Earthquake Research Institute).

れる。今回解析に用いた観測点は、おおそ数 10 km 間隔で設置されている気象庁東京管区気象台管内の地震観測点と、気象庁に一元的に収集されている東京大学地震研究所、名古屋大学、京都大学防災研究所、科学技術庁防災科学技術研究所、工業技術院地質調査所、東京都、神奈川県温泉地学研究所の上下動短周期地震計のリアルタイムデータ (248 地点) である。これらを合わせると平均の観測点間隔は約 20 km となる。

Fig. 4 左は実際の観測データから作成した 32 秒間のビットパターンの一部である。このような観測された地震波形の観測パターンは要素が 1 と 0 しかないベクトル $\mathbf{X}$,

$$\mathbf{X} = (x_1, x_2, x_3, \dots, x_n) \quad (1)$$

として表記することができる。このベクトルの要素数 n は、計算に用いる秒数と観測点数との乗数になる。

一方、震源のある場所に仮定して理論走時計算を行うことにより、P 波到着前を「0」、到着後を「1」とすれば同様のビットパターンを作成することが可能である (Fig. 4 右)。今、 N 個の震源をあらかじめ仮定し、震源 i に対するモデルパターンを、

$$\mathbf{M}_i = (m_{i1}, m_{i2}, m_{i3}, \dots, m_{in}) \quad (i = 1, 2, 3, \dots, N) \quad (2)$$

Fig. 3. Example of making bit pattern from 100 Hz sampling seismic waveform data. Each lines show bit pattern, trigger status, waveform, STA (Short Term Average) and LTA (Long Term Average), respectively.

と書くことにすると、 N 個のうちのどのモデルパターンに地震波形の観測パターンが最も類似しているかを探すことは震源決定を順問題として解くことに他ならない。

このような二つ、あるいは複数のベクトルの類似度を比較するためにここでは最短距離法とよばれる手法を用いた。最短距離法は、二つ以上のパターン $\mathbf{M}_i$ に対して例えば、

$$\|\mathbf{X} - \mathbf{M}_i\| = \left\{ \sum_{k=1}^n (x_k - m_{ik})^2 \right\}^{\frac{1}{2}} \quad (3)$$

を算出し、準備したモデルパターンのうち最小の値をもつ観測パターンを識別する手法である。なおパターン認識に関しては、例えば、舟久保 (1991) による解説がある。

前述のように、観測される振幅変化と理論到着時を表わすパターンは「1」「0」とに 2 値化されている。C 言語ではデータ間のビット演算が容易に行えるため、実際には、(3) 式のような加減乗除算を行わなくても次式の、

$$\overline{\mathbf{X} \oplus \mathbf{M}_i} \quad (\text{XOR の補数表現}) \quad (4)$$

ビット演算 (排他論理和の補数を求める演算) を毎秒実行すればよい (Fig. 5)。これは計算時間短縮に大きな効果がある。ここで補数を取っているのは、観測パターンとモデルパターンが合致すればするほど 1 が増えるという点数化を行うためである。式 (4) は観測パターンとモデルパターンの振幅変化がともに閾値を越えたとき、あるいはともに越えなかったときに 1、そうでなければ 0 が得られることを意味している。この 1 と 0 を加算することによって、二つのパターンの合致度 (以下スコアと言う) を求め、対象としている時刻にどの領域で地震が発生した可能性が高いかを毎秒識別することになる。実際の計算では地震をより識別しやすいように、震央位置に近い先着の 3 観測点、また P 波の到着時、S 波の到

[a] OBSERVED

[b] MODEL

Fig. 4. Example of observed bit pattern (left) and model bit pattern calculated for an assumed hypocenter for 32 s (right).

Fig. 5. An example calculation of “exclusive or (XOR; $X \oplus M_i$)” and “exclusive or complement ($\overline{X \oplus M_i}$)”. Open squares show 0 (=bit off) and close squares show 1 (=bit on).

着時付近などでパターンが合致した場合、通常よりも多くスコアを与える重み付けも試みた。

仮定する震源の個数や用いる観測点数は計算機的能力により制限されるが、今回は Fig. 6 の●印に示すような格子点上（緯度経度方向に 0.25 度間隔、深さ方向には 10 km と 60 km の 2 層）に仮想震源を置き、気象庁で

Fig. 6. Virtual hypocenters (circles) and the seismic stations (crosses) used in this study.

用いている標準的な走時表 [浜田 (1984)] から 248 地点の観測点 (+印) までの走時を計算した。モデルパター

Fig. 7. Spatial distribution snapshot of the score. A star shows epicenter determined by manual.

ンの総数は1980個である。モデルパターンと観測パターンを比較する時間区間は試行錯誤の結果4秒としている。この設定はマグニチュードで言えば内陸部でおおよそ2.0である。

地震が発生した場合に高いスコアが得られるのは、実際の震源付近にモデルパターンで仮定した震源がある場

Fig. 8. Change of the score with time. Time 0 (dashed line) shows origin time of the event.

合である。これらの仮定した震源要素をスコア順に並び、上位9個の緯度、経度の平均と分散を求める。分散が0.3以下で平均値が求められた領域で地震が発生したものと識別する。Fig. 7は本アルゴリズムでの識別結果をスコアの空間分布で示した例（☆印は手で検出して震源決定を行った結果）、Fig. 8は同様に震央におけるスコアの時間変化を示した例である。震央の位置、地震の発生時刻ともに正しく識別されていることが分かる。

§ 3. 地震識別結果と考察

Fig. 9は2000年1月11日から2月17日までの約1ヶ月間（途中、電源切り換えに伴う数日の中断を含

Fig. 9. (a) Epicenters of successful discrimination. (b) Epicenters of false discrimination.

Fig. 10. Relationship between the number and magnitude of events discriminated successfully by the method in this study. (a) events occurred in inland area, (b) events occurred in shore area and deep region (Dep. > 100 km).

Fig. 11. Comparison between numbers of events discriminated by "Group trigger" method and the method in this study. ■ shows number of discriminated events without corresponding earthquakes. ■ shows number of events discriminated successfully.

む), 本手法によるプログラムを実行させた結果である。(a), (b)それぞれでプロットされている地震はこの期間中に気象庁によって手動で震源決定された震央である。(a)は本アルゴリズムで地震識別が成功した地震, (b)は本アルゴリズムで地震識別に失敗した地震, つまり識別自体できなかった地震と震源位置を誤って求めた地震の震央である。なおここでは気象庁の手動による検測後に震源決定した地震の発生時刻および震央と, 本アルゴリズムの地震識別に基づくそれらとを比較し, 地震の発生

時刻が±2分以内, 震央の決まり方が内陸の地震では緯度経度方向とも±1度以内, 沿岸海域では±2度以内のものを地震の識別が成功したものと考えた。Fig. 10は内陸の地震と海域・深発地震(深さ>100 km)に分けて識別成功の個数をマグニチュード別に示したものである。

Fig. 9とFig. 10に示すとおり内陸ではマグニチュード2.5以上, 沿岸海域ではマグニチュード3.5以上の地震はほとんど取りこぼしなく識別できている。ただし, 内陸で発生する地震の場合, Fig. 10(a)を見るとあらかじめ識別が可能であろうと想定していたマグニチュード2.0からマグニチュード2.5までの地震で約3割の地震が取りこぼされている。取りこぼされた地震の多くは房総半島中央部で発生した地震である。ここでは地殻に強い減衰域があるとされており[井出(2000)], この領域を通る地震波の振幅の減少が著しく, マグニチュード2.0程度の小さな地震では観測される地震のパターンと理論パターンがそれほど一致しなかったためと考えられる。このように地域的に強い減衰や大きな走時異常がある場合, 当然のことながら今回のように標準的な走時表から単純に作成したパターンでは識別が困難になる。これとは別に, 設定した4秒間では最初に地震波形が捉えられた観測点以外が揺れない(1観測点のみしか揺れない)ため, 本質的に識別がむずかしい領域もある。なおFig. 9(b)に栃木県西部でやや大きな規模の地震の識別に失敗した例があるが, これは前の地震との発生間隔が9秒と短く, 地震が短期間に続発して識別できなかった例である。ちなみに今回識別に成功した最も小さな地震は栃木県西部で発生したマグニチュード0.6の地震である。

Fig. 11は従来のグループトリガ方式で地震を識別した処理数と本方式で地震識別を行った処理数を比べたも

Fig. 12. Comparison among numbers of hypocenters determined by manual, the method in this study, and two present automatic determination methods: one uses only P phase data and another uses both P and S phases data picked up automatically.

のである。濃い網掛けの部分、識別したイベントのうち実際に地震が存在したものである。一見、本方式による処理数が減っているように見えるが、これは識別できる地震のマグニチュードの下限が上がっているためである。マグニチュードの小さな地震まで識別可能にしようとする「地震と識別したが地震は存在しなかった」というケースが増えるというトレードオフの関係が見られる。

Fig. 12 は左から順に、手動による震源決定処理数と、本方式、および現在気象庁で用いられている 1 次自動処理 (P の自動検測値のみで行う震源決定)、2 次自動検測 (P, S の自動検測値を用いる震源決定) の地震識別処理数、震源決定数を比較したものである。グラフでは「地震は存在したが地震識別、あるいは自動震源決定処理を行わなかった数」「地震は存在し、地震識別、あるいは自動震源決定の処理を行ったが結果が不適切だった数」「地震は存在し、地震識別、あるいは自動震源決定処理を行って、かつ結果も良好だった数」「地震が存在しないのに地震識別、あるいは自動震源決定処理を行い、実際には存在しない地震をつくり出してしまった数」をそれぞれ別の模様で示している (存在しない地震については負の数として示している)。なおこの場合の自動震源決定とは波形から自動検測を行い、通常の震源決定作業で求

められたものを言う。また地震識別および自動震源決定が成功したか否かの判断は、前述のとおり、人間が手動で検測し震源決定をした地震と対応できるか否かで行っている。

ここで注目すべき点は「地震が存在しないのに誤って地震識別、あるいは自動震源決定処理を行い、実際には存在しない地震をつくり出してしまった数」の減少である。これは当初の目論見どおり、トリガ検出の時点でトリガされた観測点の出現順序や組み合わせが見かけ速度を満たさなかった場合、その後の処理を中断できたことを示す。厳密に言えば地震識別の数とその後のプロセスで処理される震源決定の地震数を直接比較するのは別次元の話であるが、トリガ検出から震源決定までの一連の処理で本方式を従来のものと併用すれば、誤認の地震を大幅に減らすことができると考える。

なお本文では処理速度に関してはあまり触れなかった。現在気象庁で行われている手法で自動検測をするためには 1 観測点当たり 1 秒程度の時間が必要であり、たくさんの観測点を検測しなければならない大きな地震ほど処理時間 (3 分程度) がかかる原因となっている。本方式では、震央精度はさほど高くないものの、わずかの時間 (数秒以内) におおよその震央位置を求めることができる。これは津波警報発表のような緊急に情報が必要

な作業には極めて有効である。また一般に観測網が高密度化すると作業量、処理量は非常に増える。今後地震計の低価格化、多目的化に伴って観測網がさらに高密度化することが予想されるが、そのときには処理の高速性と震央の高い識別能力という点で本方式は大いに威力を発揮するであろう。

§ 4. まとめと今後の課題

「確かな地震識別」、「即時震源決定」という観点から見た時、本手法は簡易かつ有効な処理方法と考えられる。さらに、今回行ったような非常に単純なパラメータチューニングでも内陸ではマグニチュード 2.5 以上、沿岸海域ではマグニチュード 3.5 以上の地震に対してその機能は實際上満足できる結果が得られた。従来自動処理に関しては、地震波の相をいかにして正確に検出するかということ、インバージョンをいかに安定して解くかということに努力が払われてきたが、今後は本論文で示したアルゴリズムと従来の手法とを組み合わせ、本来の意味での「自動」、つまり一連の作業に人間の介入が全くなく、かつ信用できる結果が得られるという自動処理システム開発が求められていくであろう。

今後の残された課題としては、

- ①現状は観測波形からパターン作成をする際に振幅の STA/LTA を用いているが、AR 係数等の時間変化に注目し、より地震波の特徴を表わすパラメータのパターン化を試みること、
  - ②識別したい地震、識別したい領域用にパラメータチューニングを詳細に行うこと、
  - ③モデルパターンを作成する際、3 次元的な減衰や走時の地域特性を考慮した走時表を用いること、
  - ④異なる場所ではほぼ同時に地震が発生した場合や、地震が頻発する群発地震の場合でも対応できるように改良すること
- などがある。

謝 辞

本解析には気象庁に一元的に収集されている東京大学地震研究所、名古屋大学、京都大学防災研究所、科学技術庁防災科学技術研究所、工業技術院地質調査所、東京都、神奈川県温泉地学研究所のリアルタイム波形データを使用させていただきました。Fig. 2 は東京大学地震研究所地震地殻変動観測センターから提供していただいたものを使用しました。

本論文をまとめるにあたり、気象庁地震火山部の松森敏幸氏、上野 寛氏、気象研究所の吉田康宏氏、東京大学地震研究所の卜部 卓氏、酒井慎一氏には有益な議論をしていただきました。また本論文は高波鐵夫氏と村井芳夫氏の適切な査読により改善されました。記して感謝いたします。

文 献

- 舟久保 登, 1991, パターン認識, 共立出版, 172 pp.
 浜田信生, 1984, 近地地震用走時表の再検討, 気象研究所研究報告, **35**, 109-167.
 堀内茂木・松澤 暢・長谷川 昭, 1999, 地震が多発した場合にも処理可能な地震波自動処理システムの開発, 地震 2, **52**, 241-254.
 井出 哲, 2000, 震源パラメーターと地殻構造情報の同時推定とその群発地震・広域データへの適用, 日本地震学会講演予稿集 2000 年度秋季大会, P066.
 前田直樹, 1985, 地震波自動処理システムにおける読み取りおよび評価, 地震 2, **38**, 365-379.
 松村正三, 1989, コンピュータによる地震観測のシステム化とその変遷, 地震 2, **42**, 371-390.
 森田裕一・浜口博之, 1984, 自己回帰過程による地震波初動の自動検出とその信頼区間, **37**, 281-293.
 白井克彦・徳弘一路, 1979, 地震波初動の自動検出, 地震 2, **32**, 141-147.
 TAKANAMI, T. and G. KITAGAWA, 1988, A new efficient procedure for the estimation of onset times of seismic waves, J. Phys. Earth, **36**, 267-290.
 卜部 卓・東田進也, 1992, win—微小地震観測網波形検出支援のためのワークステーション・プログラム(強化版), 地震学会講演予稿集, no. 2, 331.
 横田 崇・周 勝奎・溝上 恵・中村 功, 1981, 地震波データの自動検出方式とオンライン処理システムにおける稼働実験, 地震研究所彙報, **55**, 449-484.