



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

THÈSES CANADIENNES

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

THE UNIVERSITY OF ALBERTA

PATTERN ANALYSIS OF SEISMIC RECORDS: AN AUTOMATED APPROACH

by

(C) LAWRENCE HUAN TRONG LE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN

GEOPHYSICS

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA

SPRING, 1987

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights; and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-37774-7

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR

LAWRENCE HUAN TRONG LE

TITLE OF THESIS

PATTERN ANALYSIS OF SEISMIC RECORDS:

AN AUTOMATED APPROACH

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED SPRING, 1987

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED)

Lawrence Huan Trong Le

PERMANENT ADDRESS:

R. R. 2

PORTLAND, Ontario

K0G 1V0

DATEDFebruary 3.....19 87

THE UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled PATTERN ANALYSIS OF SEISMIC RECORDS: AN AUTOMATED APPROACH submitted by LAWRENCE HUAN TRONG LE in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE in GEOPHYSICS.

.....*[Signature]*.....

Supervisor

.....*[Signature]*.....

.....*[Signature]*.....

.....*[Signature]*.....

Date.....January 29, 1987.....

DEDICATION

To

the HOPE of
my FAMILY REUNION

ABSTRACT

Geophysical interpretation of seismic records involves recognition of major coherent seismic events and translation of these major events to subsurface geological structure. An automated procedure for seismic record analysis has been developed for this process. The procedure involves three steps. Initially, traces are represented as strings of local peaks with features. The first step involves trace-to-trace correlation using a string-to-string matching algorithm in which peaks are matched by a least cost criterion. Each correlation pair is assigned a correlation confidence. In the second step, parallel analysis corrects mis-matches by considering the general trend around the anomalous pairs. Matching pairs are then connected to form coherent events. Each coherent event is also assigned a confidence which reflects its reliability and importance. In the last step, a shot record is divided into zones characterized by complex attributes and other discriminating properties.

The scheme has been shown to be a promising tool for picking major events in a test using common shot gathers acquired from a survey in Saskatchewan, Canada. Most of the major coherent events detected by the scheme are consistent with those recognized visually. Major reflections correspond to events of high confidence and the seismic zones show good correlation between shot records.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. Edo Nyland for his advice, criticism and guidance throughout the research and preparation of this thesis. I am also grateful to Mr. Xuanzhi Wu for his tireless guidance and suggestions in my research. From many discussion between both of us, I have profited a lot. Further thanks are due to the members of my examination committee, Dr. E. Kanasewich, Dr. S. Sheinin and Dr. T. Marsland for their helpful comments.

I am indebted to Phyllis Tripp and Muriel Tait who have been helping me with zeroxing and letter-typing.

It is gratefully acknowledged that this thesis has been made possible with the financial support from the Natural Sciences and Engineering Research Council of Canada, the University of Alberta and the Canadian Society of Exploration Geophysicists in the form of scholarship and fellowship.

I can hardly express my heartfelt gratitude to my Canadian parants, Mr. Marshall and Doris Crowe, whose love, encouragement and generosity have supported me to continue my education.

Last but not least, I am deeply grateful to Ping, whose contribution in terms of encouragement, patience and understanding is immeasurable.

Chapter	Table of Contents	Page
1.	INTRODUCTION	1
1.1	Significance of Artificial Intelligence Techniques in Geophysics	8
1.2	An Automated Approach in Seismic Interpretation ..	11
2.	BACKGROUND KNOWLEDGE	16
2.1	Introduction	16
2.2	Statistical Pattern Recognition	18
2.3	Syntactic Pattern Recognition	19
2.4	Basic Definitions	20
2.5	Complex Signal analysis	28
3.	SEISMIC DISCRIMINATION	35
3.1	Applications of Statistical Pattern Recognition ..	35
3.2	Applications of Syntactic Pattern Recognition ...	50
4.	An Automated Seismic Event Picking Procedure	66
4.1	Primitives and features extraction	66
4.2	Record analysis	68
4.2.1	Trace-to-Trace Correlation	69
4.2.2	Coherence analysis	77
4.2.3	Zoning	81
4.3	An illustrative example	83
5.	DISCUSSION AND CONCLUSION	113
	BIBLIOGRAPHY	116

List of Tables

Table		page
1	Zones' features of shot 45.....	106
2	Zones' features of shot 46.....	107
3	Zones' features of shot 47.....	108

Figure	List of Figures	Page
1	Seismic processing and analysis is a two way interactive and iterative closed loop system.	6
2	A common shot gather a. geometry and b. data display. The dashed curve shows the locus of the reflections from the given horizontal reflecting plane (modified from Robinson, 1985).	12
3	A block diagram for a simple recognition system.	17
4	$D(i,j)$ is the distance between the substrings A_i and B_j	23
5	Three possible paths to derive $D(i,j)$ from its neighbors.	25
6	An example of optimally transforming pattern F to E.	26
7	a. $x(t)$ with $t_d=0.48s$ and its quadrature component b. envelope c. instantaneous frequency.	30
8	a. $x(t)$ with $t_d=0.52s$ and its quadrature component b. envelope c. instantaneous frequency.	31
9	a. $x(t)$ with $t_d=0.58s$ and its quadrature component b. envelope c. instantaneous frequency.	32
10	a. Traces were flattened at 1.41 msec and filtered to 6-58 Hz passband, b. classification result obtained by instantaneous frequency and the 5-component (modified from Hagen, 1982).	38
11	Discriminant analysis to distinguish the lithology of (a) synthetic model E from model F, (b) area X from area Y (modified from Sinval et al., 1983).	41
12	Percentage contribution of each variable for the	

	case of (c) synthetic data and (d) field data (Sinvhal et al., 1983).....	43
13	a. The traces of 12 seismic coda events in Eurasia as recorded by NORSAR, b. the estimated third-order autoregressive coefficients, c. the estimated third-order features for the coda of 45 earthquakes and 40 explosions (modified from Tjøtheim, 1977).	45
14	a. Cartesian cluster graph, b. migrated seismic section (modified from Bois, 1980).	48
15	a. cluster graph for sector S2 in b, b. cluster graph for sector S3 in b (modified from Bois, 1980). .	49
16	a. A stacked section and b. result of pattern recognition (modified from Lu, 1982).....	55
17	a. Relative amplitude seismogram of bright spots. b. Scattering diagram. c. Classification result (modified from Huang and Fu, 1984).	58
18	Tree structure generated by affinity algorithm based on a. average amplitude criterion and b. average slope criterion (Gaby and Anderson, 1984). ...	62
19	Resulting morphologies based on the average amplitude criterion. A. Results obtained from maximum depth: a. depth of 3; b. depth of 5; c. depth of 6. B. Results obtained from maximum segment lengths: a. length of 18 sec; b. length of 12 sec (Gaby and Anderson, 1984).	63
20	Resulting morphologies based on the slope criterion. A. Results obtained from maximum depth: a. depth of 1; b. depth of 4; c. depth of 6. B. Results obtained from maximum segment lengths: a. length of 18 sec; b. length of 12 sec (Gaby and Anderson, 1984).	64
21	A cycle within a trace.	67
22	The non-overlapping part of two waveforms.	75

23	Location map of the area under study (Gendzwill, 1978).	84
24	Plan map showing two seismic lines (A & B) surveyed by Geophysics class 528 (modified from Gendzwill, 1978).	85
25	Seismic shot gather of shot 45 in which 1=Blairmore; 2=Souris River; 3=Prairie Evaporite; 4=Winnipegosis; 5=Cambrian; 6=Precambrian.	86
26	Seismic shot gather of shot 46 in which 1=Blairmore; 2=Souris River; 3=Prairie Evaporite; 4=Winnipegosis; 5=Cambrian; 6=Precambrian.	87
27	Seismic shot gather of shot 47 in which 1=Blairmore; 2=Souris River; 3=Prairie Evaporite; 4=Winnipegosis; 5=Cambrian; 6=Precambrian.	88
28	Result of string-to-string matching of two neighboring traces.	90
29	Result of trace-to-trace correlation of record 45. ..	91
30	A comparison of Lu's scheme (A) with mine (B).	92
31	Display shows the coherent segments in record 45 with coherence confidence.	94
32	Display shows the correlation segments in record 45 with correlation confidence.	95
33	Correlation result of record 45 after parallel analysis has been applied.	96
34	Display shows the coherent segments in record 45 after parallel analysis has been applied.	97
35	Correlation result of record 46 after parallel analysis has been applied.	99

- 36 Display shows the coherent segments in record 46
after parallel analysis has been applied.100
- 37 Correlation result of record 47 after parallel
analysis has been applied.101
- 38 Display shows the coherent segments in record 47
after parallel analysis has been applied.102
- 39 Cumulative distribution of coherence confidence
versus percentage of the segments detected in record
45.104
- 40 Curves of zone features of record 45: 1=length;
2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope. ..109
- 41 Curves of zone features of record 46: 1=length;
2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope. ..110
- 42 Curves of zone features of record 47: 1=length;
2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope. ..111

1. INTRODUCTION

During the past ten years, substantial achievements have been made in the area of geophysical data acquisition, data processing and seismic inversion (Schultz, 1985). The reliability and resolution of information available to interpreters have increased; good quality control and high multiple coverage in field recording have tremendously increased signal-to-noise ratio. New schemes in statics correction and interval velocity estimation and consideration of higher order terms in migration have improved the image quality of seismic sections. Much computing power is available to do the mathematical transformations, thus yielding more ways to represent the same data set in terms of different groups of parameters. For example, a Fourier transform gives the power distribution of the frequency components and a Hilbert transform unveils the change of frequency and energy with time. These developments which are a combination of new insights and expanded computer capabilities have led us to a much better and clearer understanding of the Earth's interior.

Despite all the effort that has been spent in producing higher quality seismic sections, the function of a geophysicist remains at the same level as it was years ago. Simple routines like picking travel times, examining a velocity profile and setting up a preliminary model for migration are done manually by human interpreters. Since the

introduction of digital processing some twenty years ago, massive volumes of seismic data have been processed and analyzed, but in spite of such a large data base, computer software is not sophisticated enough to help geophysicists make even simple judgement decision, or monitor the data processing flow. Even though computer usage is more pervasive than ever, the nature of the computer programs is virtually unchanged. They are number-crunching systems, not decision-making aids.

Among these number crunching procedures, amplitude recovery is extremely important for evaluation of the sedimentary environment and delineation of lithological traps favouring the accumulation of hydrocarbon. The amplitude after processing should reflect a sequence of reflection coefficients which is a representation of lithology. Further, the mapping of a horizon is based on the consistency of amplitudes across a seismic record. Therefore, amplitude is a very important feature and demands special preservation.

In addition to the effects of noise, the amplitude of the seismic signal arriving at a geophone is affected by a number of mechanisms such as the reflection and transformation from each acoustic discontinuity, and the direct attenuation of signal amplitude due to frequency loss, and spherical divergence.

Techniques for the design of inverse gain functions to preserve relative reflection amplitude variations within and

between traces fall into statistical and deterministic categories (Lindseth, 1982). Both approaches attempt to correct traces for average attenuation rates while preserving instantaneous variations caused by changes in subsurface acoustic impedances. Deterministic approaches define general models to describe many of the possible factors affecting amplitudes. Statistical approaches such as the sliding window gain equalization, on the other hand, produce average gain functions.

The traditional approaches to gain compensation make certain assumptions, principally that the decay is smoothly exponential and the quality factor, Q , is constant within the signal's bandwidth. The model ignores the significant drop of amplitude after each major reflection. The statistical approach is also based on the assumption that the decay is smooth, though not necessarily exponential. In this technique, it is difficult to choose a proper width of the sliding window applied to the data.

Perhaps the most feasible method of gain equalization is to apply a fixed curve to correct for spherical divergence on all traces of a record, then process to eliminate some part of the noise. After this, we can recognize the major reflections and set up a simple model. The original gain compensation may then be reversed to restore the trace to its original form. A more reliable recovery of the signal's amplitude can then be made by using different gain functions between major events.

Another example of processing comes from seismic deconvolution. The ideal source signal would be a spike, but due to the nonlinear effects about the shot point, the source signal is a wavelet of finite breadth. This source wavelet is further modified by multiple reflections and absorption in the earth. Part of the wavelet's energy is also reflected upward by a sequence of reflection coefficients of the interfaces. The reflected energy is recorded by detectors and other recording instruments. All these changes on the wavelet can be visualized as a convolution. A seismic trace is thus considered as the result of convolution of a source wavelet with absorption and multiple effect, reflectivity sequence and recording response.

The reflectivity sequence represents the desired result of a seismic experiment. The other components which characterize the properties of the experiment are usually grouped together and called the seismic wavelet. Deconvolution is an inverse operation to retrieve the reflectivity sequence by removing most other effects.

Statistical deconvolution begins with the convolutional model with nothing known about the system response, except for a number of necessary assumptions regarding the statistics of the time series. One process of this type is predictive deconvolution.

Predictive deconvolution assumes that the sequence of reflection coefficients is random. Rarely is this true.

Depositional environments are the result of a complex interplay of sedimentary processes, but a vertical sequence from such an environment usually shows some cyclicity. Therefore, the assumption is invalid in most areas. Indeed, there has been a considerable trend in recent years to use whatever physical knowledge of the system can be gained to reduce the number of unknown parameters of the purely statistical model. New effective approaches in wavelet estimation and model-based deconvolution can be based on known major reflections and geological and well log data (Oldenburg et al., 1981).

Migration is an operation to reconstruct the reflector's surface from the seismic section. The process relies heavily on the local model of good subsurface velocities and the results from the operation are only as good as the model provided. A preliminary model can be set up from an automated coherent event detection algorithm. An iterative feedback of comparison between the model and the migration results can then be used to revise the initial model. The computer, in this step, might act interactively with the interpreter to monitor the operations.

In summary, seismic data processing and analysis can be thought of as a two way interactive and iterative signal processing - information extraction - model building closed loop system (see Figure 1). In order to make the loop function more efficiently and increase the speed of interpretation, the computer must take a partial role in

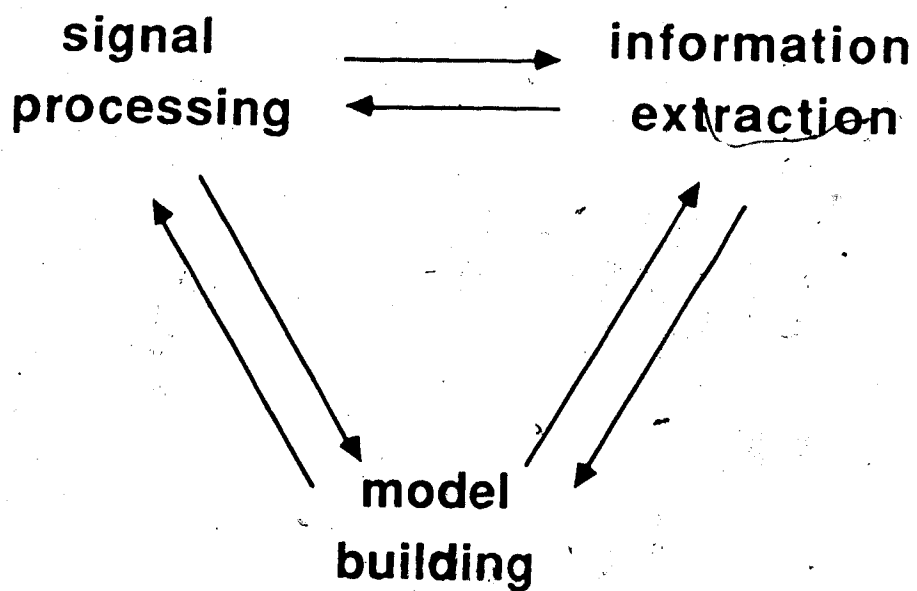


Figure 1.... Seismic processing and analysis is a two way interactive and iterative closed loop system.

decision-making and result evaluation.

Geophysical interpreters very quickly encounter data that is not easily expressed in numerical form, but is nevertheless very valid. They encounter the barrier of sharpening their understanding using empirical knowledge which may be soundly, but intuitively based. All this inherent difficulty comes from the fact that integrated induction, reasoning and judgement of contrary evidence play an essential role in geological or geophysical interpretation. Most of the judgement decisions made by interpreters are based on rules of thumb and integrated reasoning from many disciplines, and this knowledge is hard to reduce to a set of equations. The area of study between geology and geophysics is rife with examples that can be cast into this form. Although each of these problems can be considered a numerical problem, they are better thought of in the context of non-numerical programming in order to automate the geological interpretation task.

It is a common misconception and a considerable obstacle to the advancement of geophysical theory to think of a computer as solely a tool for the manipulation of numerical information. Although all geophysical information can be forced into this mould, it is better in many cases to consider the computer as a tool for manipulating logical patterns. Automating interpretation demands the computer go beyond conventional numerical processing into a nonnumeric problem-solving domain. In these problems, the computers

will deal with data sets in which the solution path is unknown and justify their reasoning when the available information is incomplete. Traditional programming fails to provide such methodology because complete information is required for a decision, but heuristic programming developed in the field of artificial intelligence and pattern recognition seems to suggest a promising tool for an automated interpretation system.

A heuristic technique is a rule-of-thumb result which can not be rigorously proven. It is a concept which, through experience, reliably guides action. For example, geological relationships are notoriously difficult to code as numerical information. They should not be coded that way if the relationships are to be used to reduce the ambiguity in interpretation often associated with geophysical data; they should be coded as lists in the sense these entities are perceived in artificial intelligence.

1.1 Significance of Artificial Intelligence Techniques in Geophysics

The use of A.I. techniques in geophysics can never replace interpreters, but can aid them in interpretation and decision-making process (Palaz, 1986). Techniques from this area aim to simulate human decision making process on computers so they can solve complex problems - problems which require intelligence if done by humans.

Some complex problems in geophysics may include locating the best drilling sites, determining the nature of the earthquakes or faults, and estimating the velocity distribution in an unexplored area. The complexity of these problems calls for an expert because expertise or experience is necessary to provide a solution. In these problems, experts deal with fuzzy data which are masked beyond recognition by other factors such as noise. The data are incomplete and uncertain.

The decisions made by experts are sometimes quite subjective and inconsistent. Even worse, due to the non-uniqueness of the solutions, several possibilities occur. Heuristic programming techniques can rank conclusions by their likelihood of being correct. The likelihood is measured by a certainty factor, which reflects the degree of confidence about the validity of a conclusion. The introduction of certainty factor gives the expert a means to check the algorithm's conclusion against his.

Moreover, the new programming techniques can transfer human expertise in given domains into effective form, so as to enable computer systems to perform convincingly and systematically as advisory consultants (Hoyle, 1986; Palaz, 1986). Knowledge representation is the heart of these A.I. systems (Waterman, 1986) and one efficient way to represent knowledge symbolically is in forms of IF-THEN production rules. An example of such rules is the criterion to detect a bright spot. As seen on a processed seismic section, a

bright spot is a reflection that is much stronger than usual for a limited distance and may indicate gases directly. If a gas reservoir really exists, then the bright spot is usually characterized by high amplitude, very low frequency and reversed polarity. In this case, a computer should be able to cope with a rule like

IF the envelope is high and
the instantaneous frequency is low and
the polarity is reversed,

THEN the feature is most likely a bright spot.

The techniques also provide the expert with the ability to communicate with the systems in terms of natural language. It is true that the communication is presently limited to technical jargon, but this is generally the way experts communicate in such circumstances. Development of such expert systems - programs that operate by reasoning from a set of facts and rules - yields knowledge of increased reliability, consistency and availability. A typical expert system, PROSPECTOR, (Waterman, 1986) developed in Stanford Research Institute, has successfully aided the geologists in locating one of the biggest molybdenum ore deposits.

Existence of knowledge-based systems can relieve a geophysicist from elementary routines like preparing the information, setting up a preliminary model, evaluating results, drawing conclusions and taking initial decision steps. He can then search for ways to extract better parameters vital for data inversion; he can also incorporate

into the results such highly significant decisions or information as only our human mind has available.

1.2 An Automated Approach in Seismic Interpretation

Seismic interpretation could be done with an expert system. It involves recognition of the major seismic events in seismic records and transformation of these events to subsurface structure. The problem requires knowledge from many fields including an interpreter's skill for picking events. Picking is marking an event on a seismic record and involves deciding what wavelets from trace to trace are from the same reflection event. The proposed automated interpretation task is best started by picking wavelets in a common shot gather, since all the information is contained in the raw data and any processing step reduces the information content of the data. Results from the initial step can then be used to construct a local model to guide the subsequent steps.

Figure 2a shows a typical geometry for recording seismic reflection data. We let s represent the horizontal coordinate of the source, r the horizontal coordinate of the detector and $q=r-s$, the offset from source to receiver. Let us suppose that the investigation is far from the source and the planar wavefront approximation is justified. The arrival time of a reflection from a given horizontal plane interface is given by

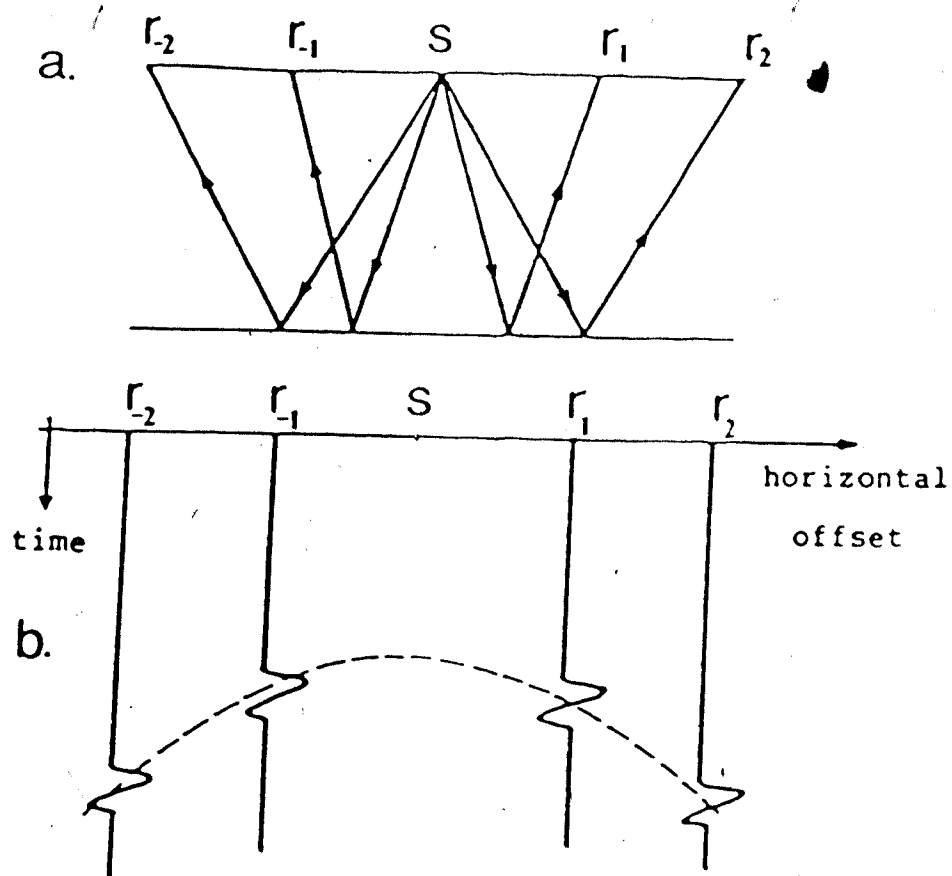


Figure 2.... A common shot gather a. geometry and b. data display. The dashed curve shows the locus of the reflections from the given horizontal reflecting plane (modified from Robinson, 1985).

$$t^2 = t_0^2 + \frac{q^2}{v^2} \quad (1.1)$$

where t_0 is the zero-offset two-way travel time and v is the (assumed) constant acoustic velocity. A reflection from a horizon is characterized by signals of high amplitude. The equation thus describes a hyperbolic trajectory of these signals in a record which displays the responses of all geophones placed side-by-side (see Figure 2b). Such a record is called a common shot gather.

Routine data processing attempts to transform such seismic data into a good subsurface image of the Earth. The performance of these techniques is always judged by the quality of the processed seismic record. The measure of this quality is usually associated with the ease of detecting the desired geological or geophysical signal in a clear background.

A collection of such signals comprises a coherent event on a record. Therefore, a coherent event characterizes a pattern and represents a set or collection of coherent signals which are associated with a geological or geophysical object such as a reflector, a refractor or a diffractor. These signals are wavelets having similar properties on neighbouring traces. Picking such coherent events is an elementary and important step in interpretation. Some examples of patterns of interest are the waveshape of a seismic trace in the neighborhood of a lithological interface, the hyperbolic patterns exhibited by

an entire set of traces in a common shot gather and the diffraction pattern due to a fault edge.

The search for an automatic picking algorithm started in the early 60's, but so far results are not very satisfactory. Some of the problems of automatic schemes (Schneider, 1971) are summarized in the following.

- a. A key mapping horizon may not be adequately picked, because of complex structure and/or poor signal-to-noise ratio.
- b. Mapping reflectors requires both time and velocity. Yet, autopicking the final section provides only the former.
- c. Autopicking schemes will tend to pick all reflections on the record section, thus presenting the interpreter with the additional burden of having to edit the reflector segment collection.

I describe below an automated procedure of seismic record analysis which will partially solve these problems. I then apply the technique to seismic reflection data.

This thesis is divided into five chapters. After discussing developments that may occur in seismic processing in this chapter, chapter two presents the necessary background for the string-to-string matching algorithm. Some discussion of the Hilbert transform is also given, since complex attributes will be used as zoning features. Chapter three reviews application of both statistical and syntactic pattern recognition in geophysics. Chapter four describes

the automated procedure and is followed by the conclusions in chapter five.

The automated procedure described in Chapter 4 is divided into three steps. Initially, traces are represented as strings of local peaks with features. The first step deals with trace-to-trace correlation employing a string-to-string matching algorithm frequently used in syntactic pattern recognition. Each correlation pair of peaks is assigned a correlation confidence. In the second step, further processing such as parallel analysis is proposed to improve the matching result. The correlation pairs are connected to form coherent events. Each coherent event is also assigned a coherence confidence which reflects its reliability and importance. Coherences with high confidence are reserved for further analysis. In the third step, a record is divided into zones characterized by complex attributes and other discriminating properties.

The idea is demonstrated with common shot gathers acquired from a survey in Saskatchewan, Canada. Most of the major coherent events detected by the scheme are consistent with those recognized visually. Major reflections are indicated by events of high confidence and the seismic zones show good correlation between shot records.

2. BACKGROUND KNOWLEDGE

2.1 Introduction

Pattern recognition involves the classification of a set of events characterized by patterns in terms of statistical features of the data or syntax information inherent in the data (Fu, 1974). Figure 3 shows a simple recognition system. The extractor retrieves information useful for classification, thus reducing the input data either to a set of statistically determined features or to a set of simple subpatterns each of which has features. These elements are usually classified according to a priori criteria which minimize some discriminant functions.

Any pattern recognition problem can be reduced to feature selection and classification. Feature selection is determined by the contribution of the features to the performance of a discrimination and recognition process. The objective of classification is to properly assign each pattern to a correct class. Many techniques of classification, such as clustering analysis (Diday and Simon, 1976; Aminzadeh and Chatterjee, 1981) and the nearest neighbor decision rule (Fu and Lu, 1977; Lu and Fu, 1978) have been suggested.

Different mathematical approaches used to solve the pattern recognition problems can be grouped into the statistical approach and the syntactic approach. Although I employ the latter approach in this thesis, the statistical

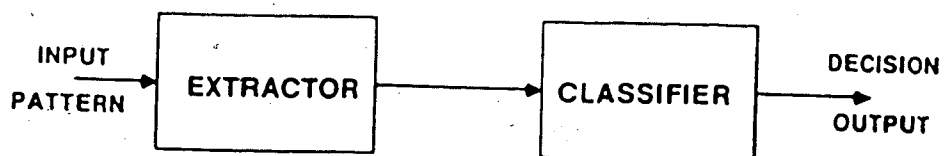


Figure 3.... A block diagram for a simple recognition system.

approach deserves some discussion.

2.2 Statistical Pattern Recognition

In statistical pattern recognition, patterns are considered to possess random qualities. Characteristic measurements or features denoted as x_1, x_2, \dots, x_n , are extracted to represent each pattern. In this sense, a pattern is represented as a feature vector $X = (x_1, x_2, \dots, x_n)$ in a N -dimensional feature space. Classification can then be done by partitioning the feature space in terms of some similarity measure into clusters or regions where each region represents a pattern class. The similarity measure quantifies the proximity or the dissimilarity of the pattern vectors and is often expressed as a metric in the feature space. The clustering results can then be used for future unknown samples.

Clearly, the selection of features and the metric has a strong influence on the results of cluster analysis. Ideal features contain no redundancies and possess great discriminating power. Measurement of an insufficient number of features will not give satisfactory classification results. The approach will become impractical and inefficient for complicated patterns with large feature sets as in picture and scene analysis. The nature of these problems requires pattern analysis and interpretation beyond mere classification. If feature sets overlap, the statistical pattern recognition approach fails and analysis

of the structural information is necessary for a more accurate classification. As a consequence, syntactic pattern recognition has been introduced (Fu, 1974; Pavlidis, 1977).

2.3 Syntactic Pattern Recognition

Unlike the statistical approach, the syntactic approach emphasizes the representation of patterns. The association of each pattern to a pattern class is completed by a description of a transformation path. The path describes how one pattern can be derived from the other in terms of some predefined transformation rules.

In the representation, a pattern is broken down into its simplest subpatterns. These simplest subpatterns are the basic building blocks composing the pattern and are called the pattern primitives. Evidently, if this approach is to be efficient, the pattern primitives should be much easier to recognize than the pattern. Moreover, the primitives should serve as basic pattern elements to provide a compact description of the pattern being analyzed. Linguistically, a pattern is represented as a string of pattern primitives $a_1 a_2 \dots a_l$. Each primitive retains its own identity by means of its own features. The string representation of a pattern is unique in the syntactic pattern recognition approach and provides a means of representing the pattern for automated recognition systems. On a shot record, a seismic trace can be loosely considered as a pattern and an example of a primitive is a peak with features such as peak amplitude,

location and duration of the waveform. A comparison between syntactic patterns is then made on their corresponding pattern strings by using the pattern feature sets. One of the classification techniques often employed is the string-to-string matching.

2.4 Basic Definitions

After the patterns have been encoded as strings of primitives, the derivation of one pattern from the other can be described in terms of step-by-step transformations. Some formal definitions of the terms that will be used in the discussion are necessary (Wagner and Fisher, 1974; Fu and Lu, 1977; Lu and Fu, 1978).

In the correlation of two strings, the primitives of one string are correlated to primitives of the other strings. In this case, the evaluation of correlation between two strings can be reduced to the correlation of their substrings. For two pattern strings, $A = a_1 a_2 \dots a_m$ and $B = b_1 b_2 \dots b_n$ where a_i and b_j are primitives and $1 \leq i \leq m$, $1 \leq j \leq n$, three types of transformations are introduced to derive one string from the other. With each transformation, there is associated a cost that measures the amount of deviation or dissimilarity between two primitives. An exact mapping in which two strings are identical yields a null cost.

a. Substitution transformation

If substring A_{i-1} matches substring B_{j-1} , A_i matches B_j if

$$a_i \longrightarrow b_j.$$

The substitution transformation, T_s has a cost $c(i, j; T_s)$.

b. Deletion transformation

If substring A_{i-1} matches substring B_j , A_i matches B_j if

$$a_i \longrightarrow \Lambda_j$$

where Λ_j is a null primitive between b_j and b_{j+1} .

The deletion transformation, T_d has a cost $c(i, j; T_d)$.

c. Insertion transformation

If substring A_i matches substring B_{j-1} , B_j matches A_i if

$$b_j \longrightarrow \Lambda_i$$

where Λ_i is a null primitive between a_i and a_{i+1} .

The insertion transformation, T_i has a cost $c(i, j; T_i)$.

For a sequence of transformations \underline{T} , which takes string A to string B , $\underline{T} = T_1, T_2, \dots, T_k$, where $T_l \in \{T_s, T_i, T_d\}$, $l = 1, 2, \dots, k$, the total cost of error transformations is given as $C(\underline{T}) = \sum_{i=1}^k c(\dots; T_i)$ where \dots represents the proper pair of matched primitives. There should be more than one

such transformation sequence. Let T be the set of all such sequences that take A to B , then the distance between A and B , $D(m,n)$, is defined as the minimum cost of total transformations

$$D(m,n) = \min\{C(\underline{T}) \mid \underline{T} \in T\}. \quad (2.1)$$

Use of distance in correlation allows the computation of the degree of mismatch between the correlated pairs, since the pairs are seldom perfect matches with one another. In reality, no two series will be identical. The distance is then a measure of the natural variation due to noise or the inherent characteristics of the system.

If we denote $D(i,j)$ (see Figure 4) as the distance between the substring of A up to i^{th} primitive, $A_i = a_1, a_2, \dots, a_i$ and the substring of B up to j^{th} primitive, $B_j = b_1, b_2, \dots, b_j$, then

$$\left. \begin{aligned} D(i,j) = \min \{ & D(i-1, j-1) + c(i, j; T_s) \\ & D(i-1, j) + c(i, j; T_d), \\ & D(i, j-1) + c(i, j; T_i) \} \end{aligned} \right\} \quad (2.2)$$

for all i, j , $1 \leq i \leq m$, $1 \leq j \leq n$, with initial condition

$$D(0,0) = 0. \quad (2.3)$$

It is seen that the choice of path is not arbitrary, but

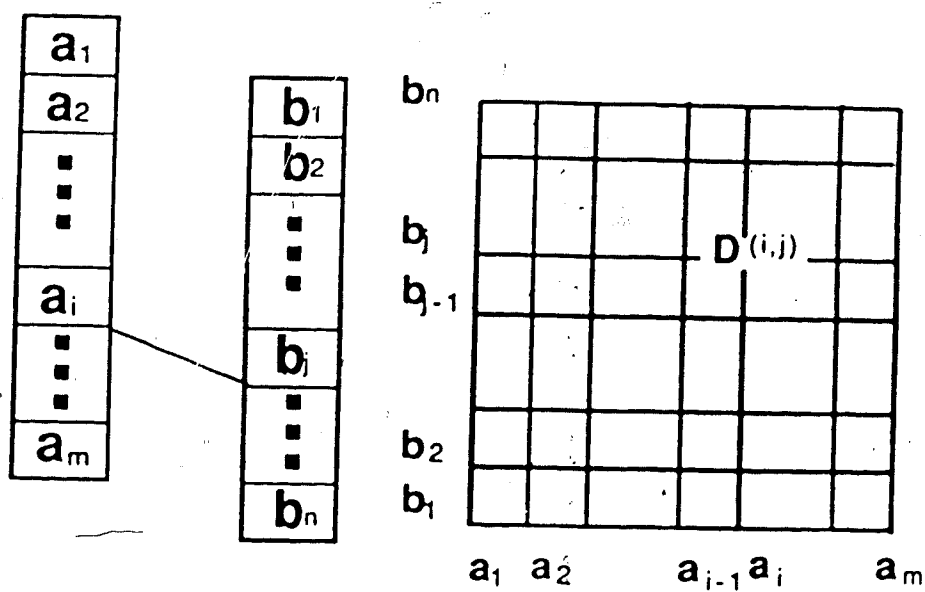


Figure 4.... $D(i, j)$ is the distance between the substrings A_i and B_j .

depends on the cost associated with each transformation as well as the accumulated cost of the subpath. The accumulation of the matching cost, $c(i,j;T_1)$ at each step becomes the total cost of the proposed match. The candidate path is then the one of least cost.

Figure 5 shows three alternative paths to derive $D(i,j)$ from its neighbors. The horizontal direction denotes deletion; the vertical, insertion; and the diagonal, substitution. The distance $D(i,j)$ corresponds to the minimum of the following:

- a. the cost of transforming substring A_{i-1} to substring B_{j-1} plus the cost of substituting a_i to b_j (Figure 5a), and
- b. the cost of transforming substring A_{i-1} to substring B_j plus the cost of deleting a_i from A (Figure 5b), and
- c. the cost of transforming substring A_i to substring B_{j-1} plus the cost of inserting b_j into A immediately after a_i (Figure 5c).

Figure 6 illustrates the path which transforms pattern F to E. Every point (i,j) of the grid is assigned a least cost $D(i,j)$ of transformation. The best path of transformation is indicated in thick dark line. The optimal transformation is

$$F = \text{acadg} \xrightarrow{T_i} \text{abcadg} \xrightarrow{T_d} \text{abcdg} | -$$

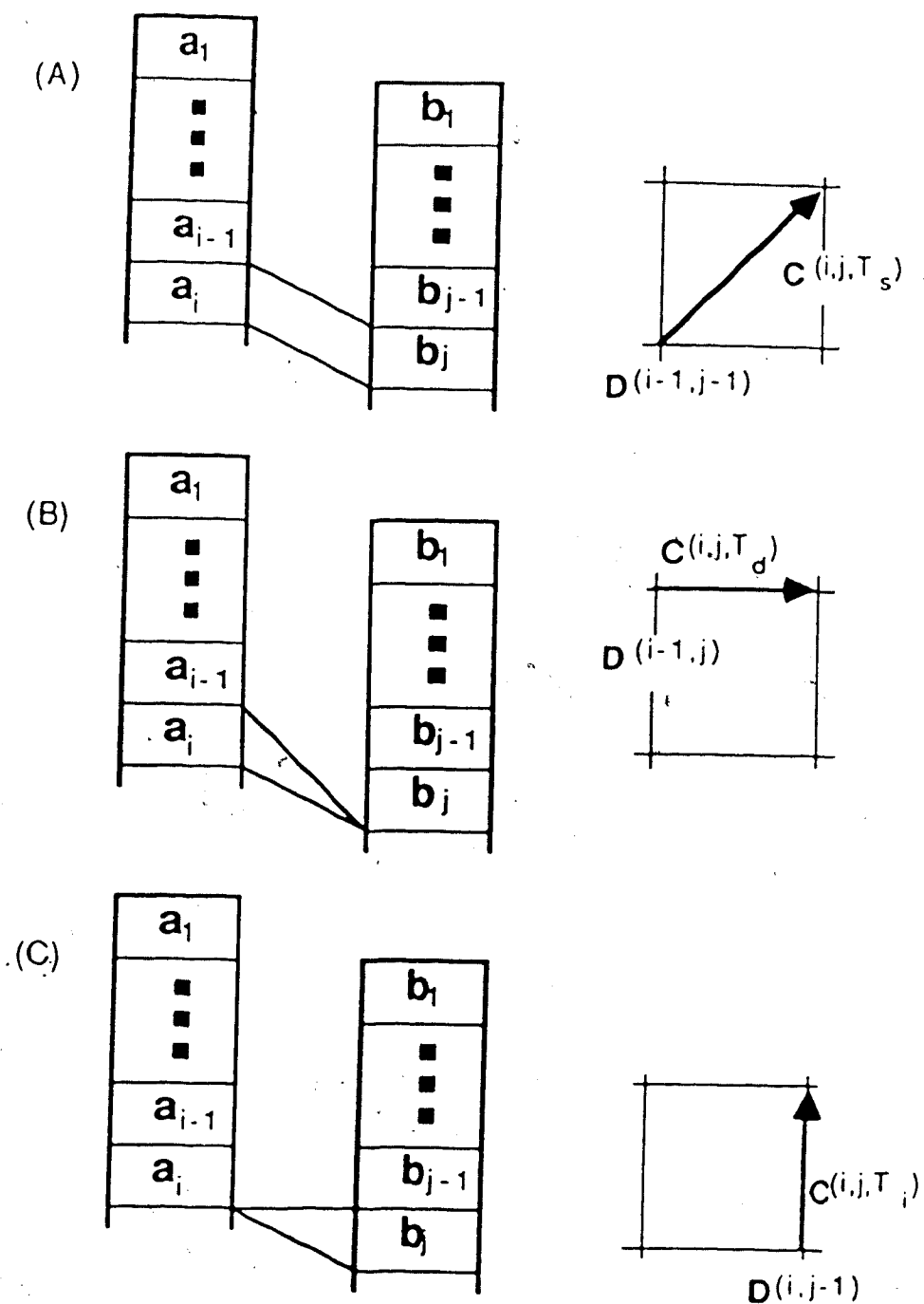


Figure 5.... Three possible paths to derive $D(i, j)$ from its neighbors.

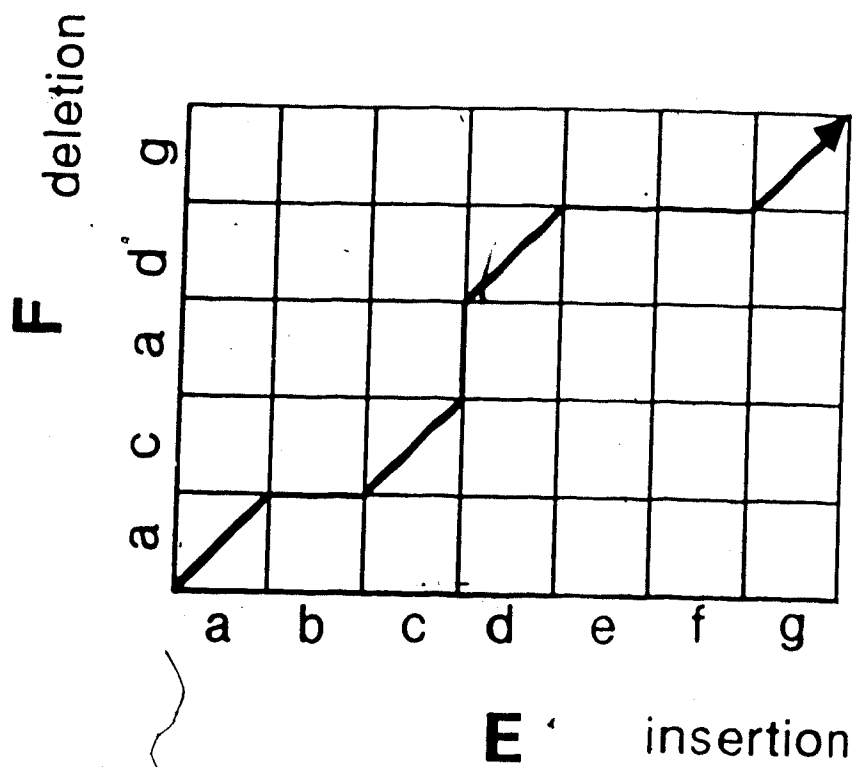


Figure 6.... An example of optimally transforming pattern F to E.

$$\begin{array}{c} T_i \\ \hline \end{array} \rightarrow abcdeg \mid \begin{array}{c} T_i \\ \hline \end{array} \rightarrow abcdefg = E.$$

If the cost of each transformation is 1, then the total distance of the mapping is

$$D(F, E) = 4.$$

The imposition of constraints is necessary to complete the global sense of the least cost path, and will be discussed in chapter 4. The selection of the transformation path according to the minimum distance criterion from its neighboring points is often called the nearest neighbor decision rule.

The string-to-string matching algorithm includes two steps. The first step is the cost computation. The costs of all three types of transformation are calculated at each point (i, j) and the distance between substrings is recorded along with the appropriate neighboring coordinate of D . The final step is just a matter of backtracking the optimal path which gives the total distance D . The optimal path shows the best correlation between strings A and B . Cost measures the degree of similarity of two patterns by comparing their features. In syntactic pattern recognition, features are not used to represent a pattern class; instead, they are used to discriminate one primitive from the other. Although feature selection is not important in a syntactic recognition scheme; a better recognition result may be obtained if

features of high discriminating power are employed in the cost calculation. In a seismic signal, complex features are useful discriminators.

2.5 Complex Signal analysis

In complex signal analysis (Bracewell, 1965; Ziemer et al., 1976; Kanasewich, 1981), a signal whose amplitude spectrum for negative frequencies does not exist is termed an analytic signal. The analytic signal representation of $\psi(t)$ of a real time function $x(t)$ is a complex signal given as

$$\psi(t) = x(t) + j\hat{x}(t) \quad (2.4)$$

where $j^2 = -1$ and $\hat{x}(t)$ is the Hilbert transform of the original signal $x(t)$ and is called the quadrature component of the complex signal ψ . In the time domain, the quadrature component is

$$\hat{x} = x * \frac{1}{\pi t}$$

or
$$\hat{x}(t) = \int_{-\infty}^{\infty} \frac{x(t')}{\pi(t-t')} dt' \quad (2.5)$$

where $*$ denotes convolution. From (2.4) or (2.5), the following complex features of the signal can be extracted:

1. envelope

$$E(t) = |\psi(t)| = \sqrt{x^2(t) + \hat{x}^2(t)} \quad (2.6)$$

2. instantaneous frequency

$$f(t) = \frac{1}{2\pi} \frac{d}{dt} \left\{ \tan^{-1} \frac{\hat{x}(t)}{x(t)} \right\}$$

$$= \frac{1}{2\pi} \frac{x(t)[d\hat{x}(t)/dt] - \hat{x}(t)[dx(t)/dt]}{\hat{x}^2(t) + x^2(t)} \quad (2.7)$$

I have made some experimental studies of the properties of the complex features using a pulse to simulate teleseismic signals. The pulse is expressed mathematically as

$$p(t) = 0 \quad t < 0$$

$$= t^2 \exp(2-8t) \sin(12\pi t) \quad t \geq 0. \quad (2.8)$$

In the study, two pulses of this form separated by time t_d are superimposed to give a composite pulse $x(t)$ where

$$x(t) = p(t) + p(t - t_d). \quad (2.9)$$

Figure 7, Figure 8 and Figure 9 show three different versions of $x(t)$, its Hilbert transform $\hat{x}(t)$, the envelope and the instantaneous frequency. They correspond to delay times of 0.48s, 0.52s and 0.58s respectively. The arrow indicates the time when the second signal arrives.

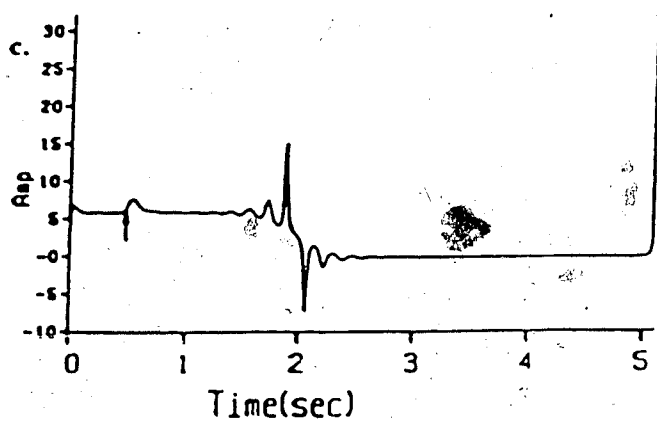
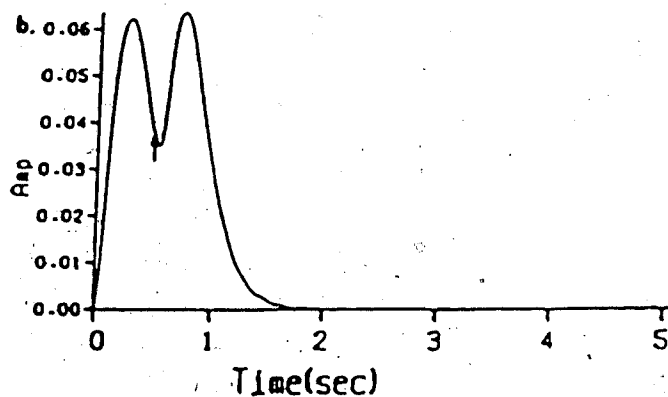
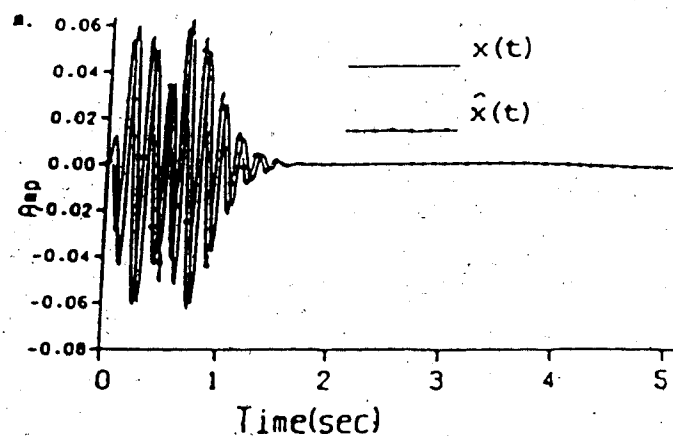


Figure 7.... a. $x(t)$ with $t_d=0.48$ s and its quadrature component b. envelope c. instantaneous frequency.

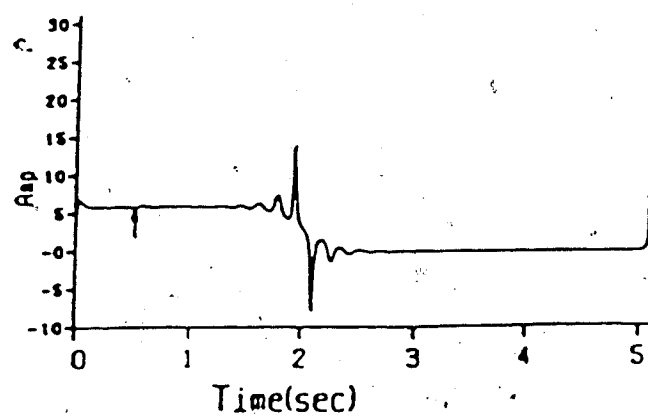
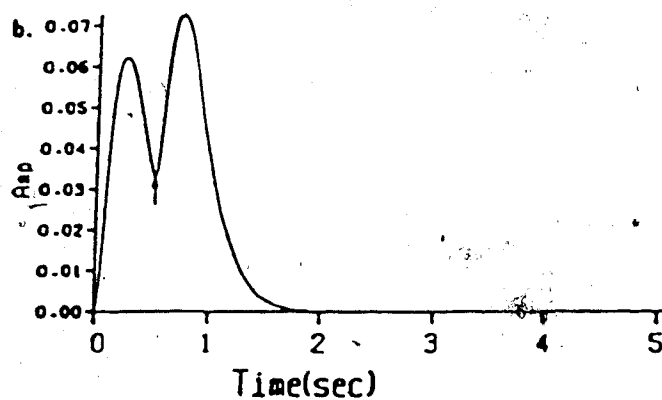
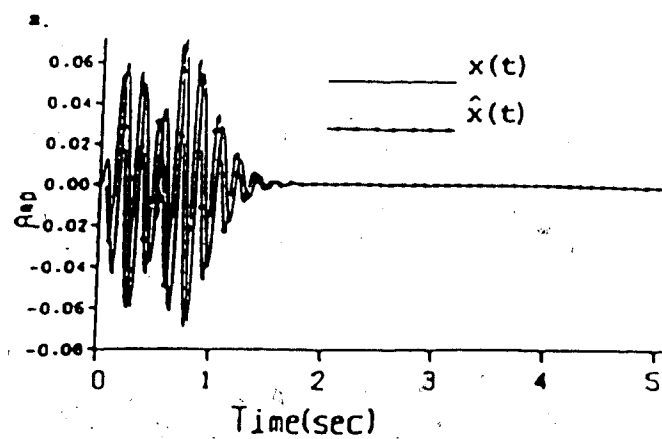


Figure 8.... a. $x(t)$ with $t_d = 0.52$ s and its quadrature component b. envelope c. instantaneous frequency.

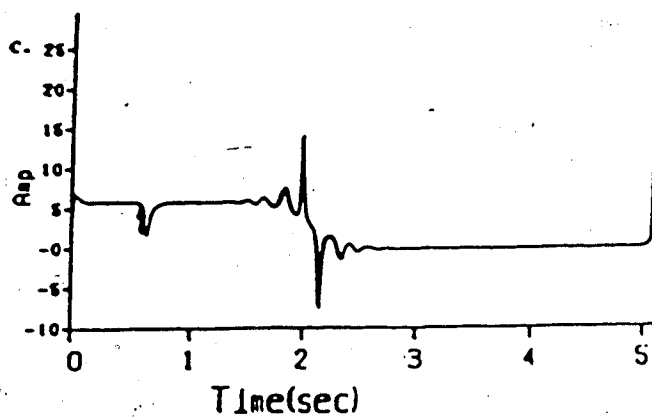
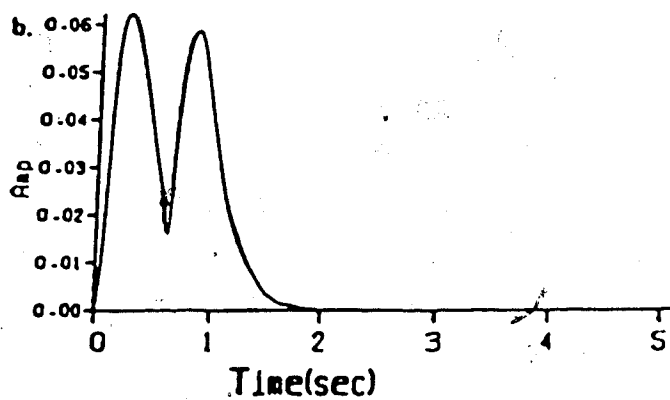
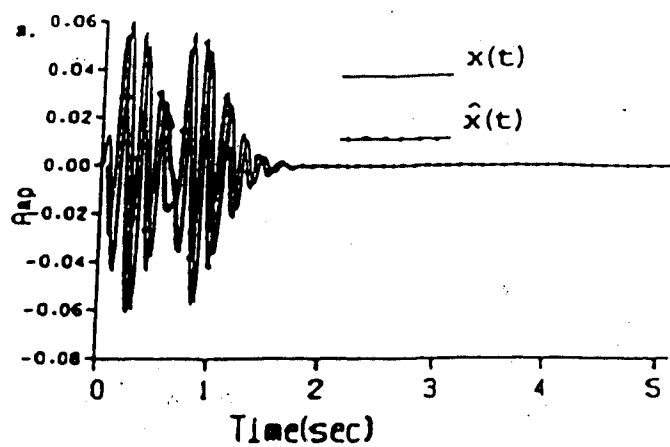


Figure 9.... a. $x(t)$ with $t_d = 0.58$ s and its quadrature component b. envelope c. instantaneous frequency.

From the computed results, two observations are made.

- a. Previous results (Robertson & Nogami, 1984; Taner et al., 1977 & 1979) show that the instantaneous frequency corresponding to the peak of a Ricker or zero-phase wavelet closely represents the central frequency of the wavelet. The concept can be relaxed to the case when the instantaneous frequency corresponding to the peak of the envelope approximates the central frequency of the artificial composite pulse, $x(t)$. Also, within a certain interval centering at the peak of the envelope, the instantaneous frequency is constant. The frequency in this example is 6 Hz and can be observed in b and c of Figure 7, 8 and 9.
- b. Kicks or sudden jumps in the instantaneous frequency plot are usually used as indicator for arrival time of the incoming energy (Aboutajdine et al., 1981; Farnbach, 1975). This observation is not always correct; however, if the signals join smoothly as indicated in Figure 8a, the kick does not exist (see Figure 8c). Also the existing kicks can be negative or positive (Figure 7c and Figure 9c).

Many techniques have been suggested to enhance the frequency of true events and suppress irregularities such as spikes (Kirlin, 1984). One of them is to weight the instantaneous frequency. Integer powers of the instantaneous envelope can be used to reduce the variance of the

instantaneous frequency and thus, reduce the amplitude of spikes (Strom, 1976). In our particular example, the weighted frequency is given as

$$f(t) = \frac{\sum_{i=-1}^{i=1} E^*(t-i)f(t-i)}{\sum_{i=-1}^{i=1} E^*(t-i)} \quad (2.10)$$

The Hilbert transform allows the separation of amplitude information from phase information of the signal. It permits extracting more parameters from seismic data. These complex parameters might contain significant information about the geological structure not easily detected by conventional analysis. Some applications of complex attributes can be found in Farnbach (1975); Taner and Sheriff (1977); Robertson and Nogami (1984). It is important to note that the complex attributes so derived represent the structural content of the signal and have true physical meaning. The instantaneous envelope reflects the strength of the acoustic impedance of the reflectors. The instantaneous frequency measured at the peak of the envelope corresponds to the central frequency of the signal calculated from the zero crossings.

3. SEISMIC DISCRIMINATION

This chapter is a review of the application of the pattern recognition techniques in seismology as well as in exploration geophysics. The papers discussed below have succeeded in demonstrating the potential and merit of pattern recognition techniques to aid seismic interpretation. The concepts or feature selection criteria applied differ vastly from one paper to another. The review is divided into two sections: statistical and syntactic.

3.1 Applications of Statistical Pattern Recognition

Hagen (1981,1982) applied principal components analysis and clustering to classify traces corresponding to porous and nonporous lithology. The statistical theory is summarized as follows. Consider a matrix X , whose N columns represent discrete time series vectors X_1, X_2, \dots, X_N of M samples each, drawn from the same population. The mean vector and the covariance matrix estimates of the population are

$$\bar{x}_i = \frac{1}{N} \sum_{k=1}^N x_{ik}, \quad (3.1)$$

$$S_{ij} = \frac{1}{N-1} \sum_{k=1}^N (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \quad (3.2)$$

where x_{ik} is the i^{th} component of the k^{th} sample vector. It is assumed that the mean and covariance contain all statistical information about the data set. Since S_{ij} is in

general nonzero, the observed sample vectors are considered to be generated from a smaller set of underlying factors which give rise to the observed dependence. An orthogonal transformation matrix A can be assumed,

$$Y = A^T X \quad (3.3)$$

such that the variance of Y given by

$$D = A^T S A \quad (3.4)$$

results in statistical independence. The columns of A are eigenvectors of S and

$$\begin{aligned} D_{ij} &= \lambda_i & \text{for } i=j \\ &= 0 & \text{for } i \neq j \end{aligned} \quad (3.5)$$

where λ_i are eigenvalues of S and are arranged such that $\lambda_1 \geq \dots \geq \lambda_N$.

The contribution of each eigenvector to the overall variance is proportional to the corresponding eigenvalue. So if the first $N < N$ eigenvalues, or principal components, provide a sufficient amount of variance measure, then each trace X_i can be estimated as a linear combination of the N eigenvectors,

$$x_{ik} \approx \bar{x}_{ik} = \sum_{n=1}^N a_{ni} y_{nk}. \quad (3.6)$$

Cluster analysis defines a distance

$$d_{AB} = \left[\sum_{n=1}^N \frac{(y_{nA} - y_{nB})^2}{\lambda_n} \right]^{1/2}. \quad (3.7)$$

as the merging criterion between two traces A and B. If d_{AB} is below a threshold value, then the lithologic properties corresponding to traces A and B are considered to be in the same class; otherwise, they are from two different classes. The probability that the group Z is in the class A is given by

$$P(Z \in A) = \frac{1/d_{ZA}}{1/d_{ZA} + 1/d_{ZB}}, \quad (3.8)$$

where B is another class.

A stacked seismic section investigated by Hagen was shown in Figure 10a. The target was a porous zone which appeared and disappeared laterally at depths from 1.66 to 1.68 seconds. The clustering used the traces 214 and 230 as cluster centers. Using the instantaneous frequency as feature, Hagen claimed that the 5-component case showed the best resemblance to the original instantaneous frequency. The classification result was shown in Figure 10b in which a dark background indicated that the lithology there was classified as porous and a light background as nonporous. The classification of trace at 195 was less certain due to

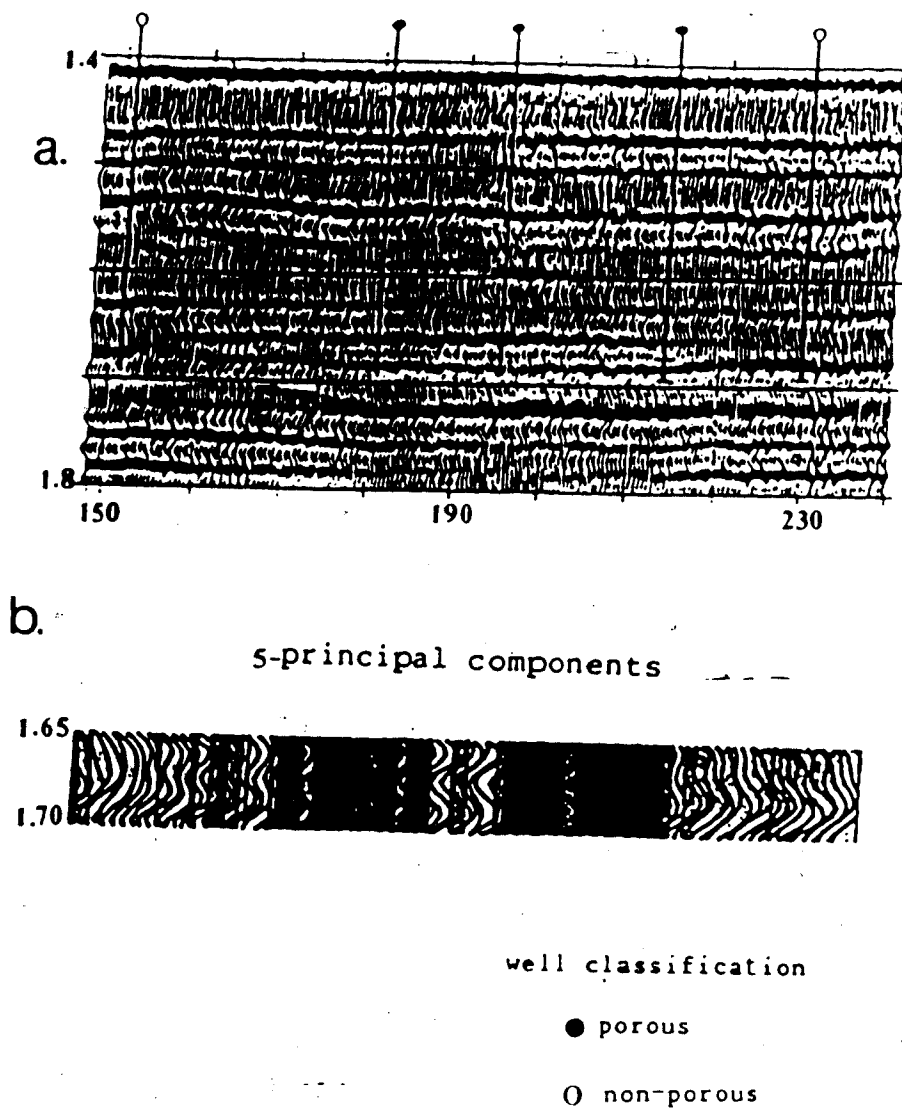


Figure 10.... a. Traces were flattened at 1.41 msec and filtered to 6-58 Hz passband, b. classification result obtained by instantaneous frequency and the 5-component (modified from Hagen, 1982).

the bad data quality in that area. In this example, data was reduced to approximately 10% of the original data base size.

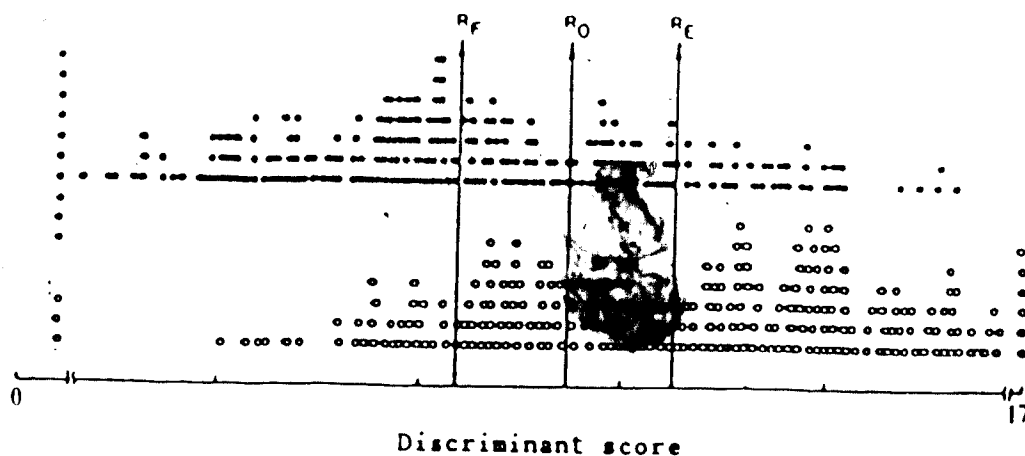
Sinvhal and Khattri (1983) established a correlation between lithology and different features abstracted from the synthetic as well as the field reflection seismograms. They used first-order Markov chains to model two different hydrocarbon-bearing formations in a sedimentary basin. One formation was sandy and the other was shaly, while both were of sand-shale-coal alternation. They suggested that Markov chains could be conveniently used to model complex processes which were subjected to influences that could not be exactly evaluated and whose changes of state could only be interpreted in terms of relative probability of occurrence. 508 synthetic seismograms and 387 real seismograms were used as training patterns. Seventeen parameters from both time and frequency domains were extracted from each seismogram and were used as features:

1. A_1/A_0 , where A_n denoted the autocorrelation function at the subscripted lag,
2. A_2/A_0 ,
3. A_3/A_0 ,
4. A_{min}/A_0 ,
5. T_1 , time of the first zero crossing,
6. T_2 , time of the second zero crossing,
7. T_3 , time of the third zero crossing,
8. T_{amin} , time of the first minimum,
9. f_M , the frequency at which maximum energy

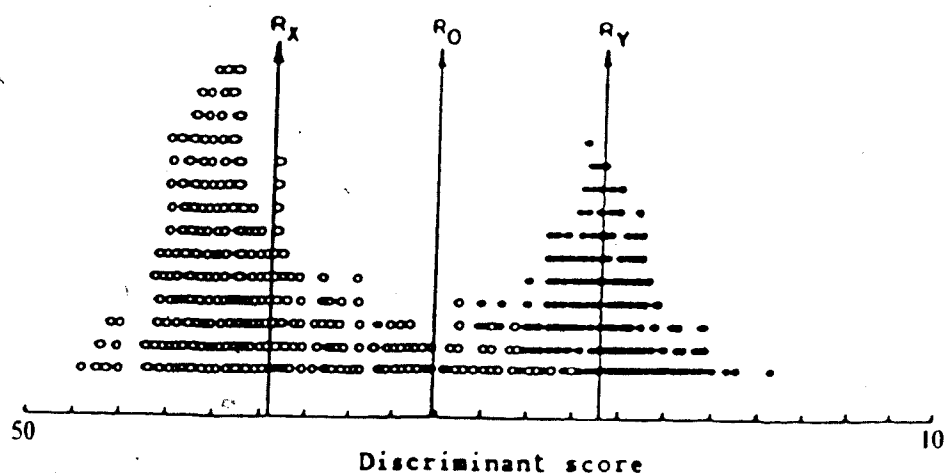
occurs,

10. f_1 , the average power weighted frequency of the power spectrum,
11. f_2 , the frequency at which the twenty-fifth percentile value of frequency weighted power occurs,
12. f_3 , the frequency at which the fiftieth percentile value of frequency weighted power occurs,
13. f_4 , the frequency at which the seventy-fifth percentile value of frequency weighted power occurs,
14. f_5 , the frequency at which the twenty-fifth percentile value of power occurs,
15. f_6 , the frequency at which the fiftieth percentile value of power occurs,
16. f_7 , the frequency at which the seventy-fifth percentile value of power occurs,
17. f_8 , the lowest frequency at which the logarithm of power decreases to half its value,

Discriminant analysis was then applied to both synthetic and real data. Each seismogram was represented as a vector in the 17-dimensional space. The vectors were then mapped to a discriminant score, R by a linear discriminant function. R_E and R_F (or R_X and R_Y) in Figure 11 were the projected discriminated scores of two mean vectors representing two different kinds of lithology. Seismograms



a. synthetic seismograms



b. field seismograms

Figure 11.... Discriminant analysis to distinguish the lithology of (a) synthetic model E from model F, (b) area X from area Y (modified from Sinval et al., 1983).

with the same value of discriminant score were plotted at different heights to avoid overlapping of points. Their mean R_0 served as a dividing line between two clusters of discriminant scores. Depending on whether the discriminant score for a particular seismogram was greater or less than R_0 , it could be classified to area X or Y. A set of 20 seismograms from areas X and Y were tested and 90% were correctly classified.

The percentage contribution of each of the seventeen variables was given in Figure 12. Positive contribution indicated that the variables were meaningful discriminators and large values corresponded to strong discriminants. Variables which gave negative contribution were detrimental and should have been discarded. This example served to demonstrate a difficulty in feature selection which was crucial in clustering analysis. Seventeen features were used in this case; nevertheless, the great number of features did not guarantee a good classification result. Some features which were thought to be good discriminators in synthetic data resulted in negative contributions on real data sets. Matlock and Asimakopulos (1986) used a similar technique to determine boundaries between different sand contents in a sand-carbonate formation.

Tjøstheim (1977) proposed the use of the prediction filter coefficients as features to discriminate seismic events. A time series $x(t)$ is stationary with zero mean and M_0 discrete samples. By adapting the time series to an M^{th}

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
1.	T_{min}	1.1	1.3
2.	T_1	6.5	0.0
3.	T_2	24.5	-0.1
4.	T_3	-3.9	0.7
5.	A_{min}/A_0	-4.0	0.0
6.	A_1/A_0	38.4	0.1
7.	A_2/A_0	2.1	5.9
8.	A_3/A_0	2.0	-2.1
9.	f_1	6.9	4.2
10.	f_2	-3.9	-0.5
11.	f_3	3.4	0.7
12.	f_4	-3.8	-1.2
13.	f_5	-1.3	-1.9
14.	f_6	-12.8	-1.5
15.	f_7	-4.4	-1.7
16.	f_8	30.5	95.3
17.	f_w	18.7	0.8

Figure 12.... Percentage contribution of each variable for the case of (c) synthetic data and (d) field data (Sinvhal et al., 1983).

order autoregressive process, an estimate \hat{x}_k of x_k , which is a linear combination of M values of $x(t)$ prior to x_k can be obtained:

$$\hat{x}_k = \sum_{j=1}^M a_j x_{k-j}, \quad k=M, M+1, \dots, M_0. \quad (3.9)$$

The estimated value of the power spectrum of the autoregressive process is given by

$$\hat{S}(f) = \frac{2\sigma_M^2 \Delta t}{|a_1 - \sum_{j=2}^M a_j \exp(-i2\pi f j \Delta t)|^2} \quad (3.10)$$

where σ_M^2 is the variance. Since the autoregressive coefficients a_j computed from the seismic traces completely determine the shape of the power spectrum, the coefficients can then be used as features for seismic events, if they are stable enough with respect to the estimation process. By assuming weak stationarity of the signal, the coefficients can be calculated by adapting the time series to an M^{th} order autoregressive process with a minimum mean-square error criterion or by using the maximum entropy principle. For short signals, a rough approximation about the power spectrum of the data can be obtained using the second- or third-order model.

Such coefficients were used as features to discriminate coda of shallow earthquakes from the nuclear explosions. Figure 13a showed the traces of 12 seismic events recorded by the Seismological Observatory Norwegian Seismic Array

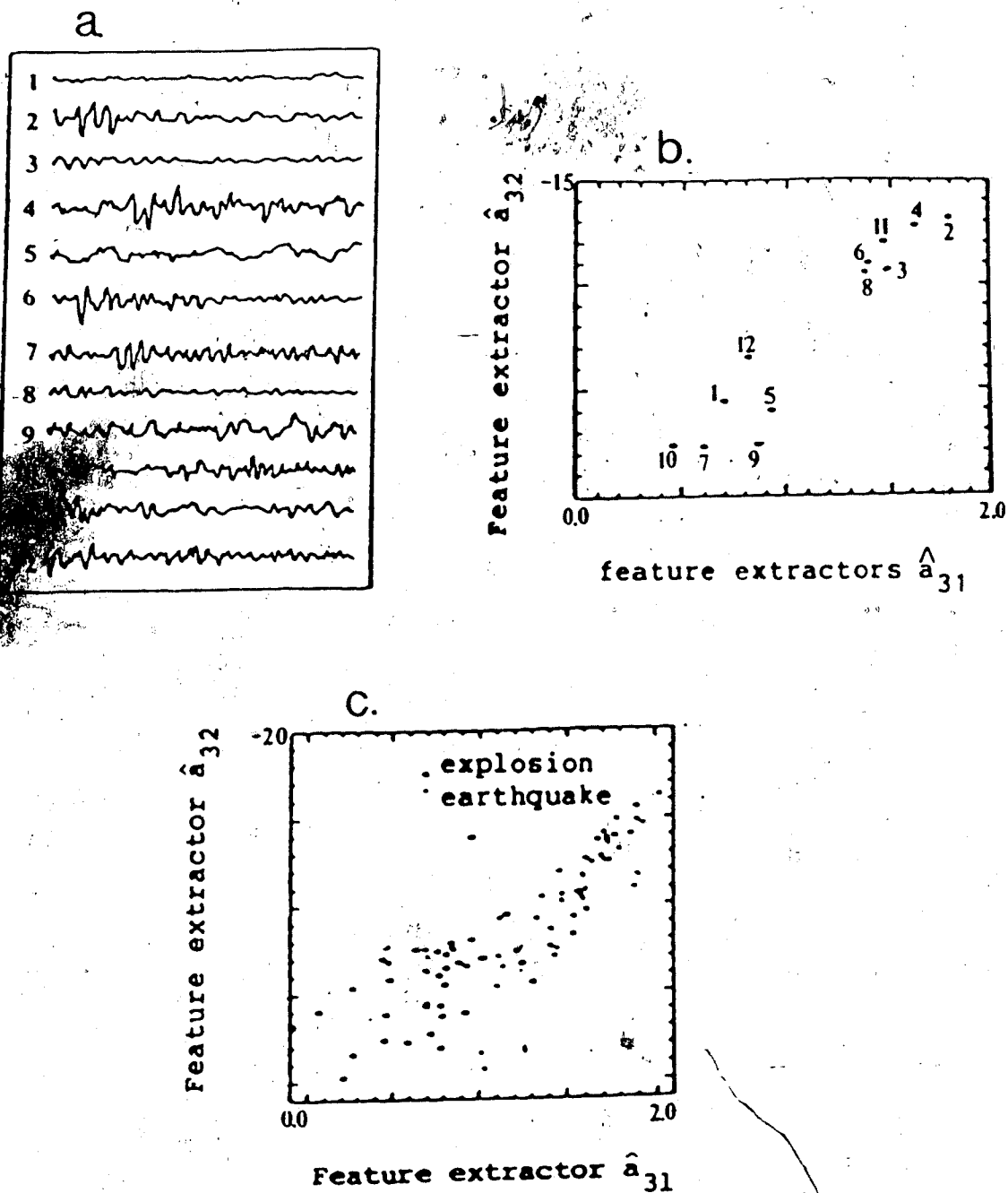


Figure 13.... a. The traces of 12 seismic coda events in Eurasia as recorded by NORSAR, b. the estimated third-order autoregressive coefficients, c. the estimated third-order features for the coda of 45 earthquakes and 40 explosions (modified from Tjøheim, 1977).

(NORSAR). Six of the traces (1,5,7,9,10,12) were from underground nuclear explosions, while the remaining six came from shallow earthquakes. All the events had their epicenters in Eurasia. Figure 13b showed the estimated third-order autoregressive feature extractors \hat{a}_{31} and \hat{a}_{32} of the 12 events. The figure showed a clear separation of two clusters representing earthquakes and explosions. However, this well-defined clustering did not hold in case where earthquakes behaved as explosion-like events as indicated in Figure 13c where 45 earthquakes and 40 explosions were studied. The difficulty could be partially resolved by the syntactic method described by Liu and Fu (1982a & b).

Bois (1980) applied the same techniques to study the lateral facies variation and the boundaries of hydrocarbon reservoirs. The autoregressive coefficients for the portion of seismic traces between the top and bottom boundaries of the reservoir are calculated. For a third-order process, each trace is specified by a dot in the cartesian feature space characterized by the coefficients a_{32}^* and a_{33}^* . When these trace sectors belong to homogeneous formations, the dots form elongated ellipses having their principal axis quite close to the main diagonal of the graph (see Figure 14a). The decision rule which formulates a criterion for cluster separation defines a pseudo-distance between two dots p_i and p_{i+1} . The distance is given by the Euclidean distance between them multiplied by the distance d separating p_{i+1} from the main axis of the ellipse formed by

the dots under study. A dot cluster is made up of dots having approximately equal pseudo-distance. A separation between two clusters occurs whenever a discontinuity appears in the successive pseudo-distance values. The boundary between clusters, then serves as indicator for the boundary of the reservoir.

The example (see Figure 14b) consisted of a seismic section 3.5 km long in an sandy series containing a bright spot in its middle. Two reservoirs were located at a depth of 1.25s and cover about 3 km. The study was to determine the boundaries of the bigger reservoir. Figure 15a corresponds to the forty traces making the sector S2. The sector belonged to the gas reservoir and extended from the 88th to the 127th traces. The excellent homogeneity of the pseudo-distances between the dots took on the form of the ellipsoid cluster. In this case, the automation indicated that there was no facies variation. Figure 15b corresponded to the sector S3. The non-homogeneity of the cluster indicated that the boundary had been met. The location of the boundary could be obtained from the trace at which the pseudo-distance had a discrete jump in value.

Bois (1981a,b,1982) used the same clustering approach to determine the nature of an unknown reservoir using the nature of a known reservoir as monitor. The process involved the analysis of 20 traces from each reservoir. If they were of the same nature, a homogeneous, elliptical cluster occurred; If they were different, two clusters were

e

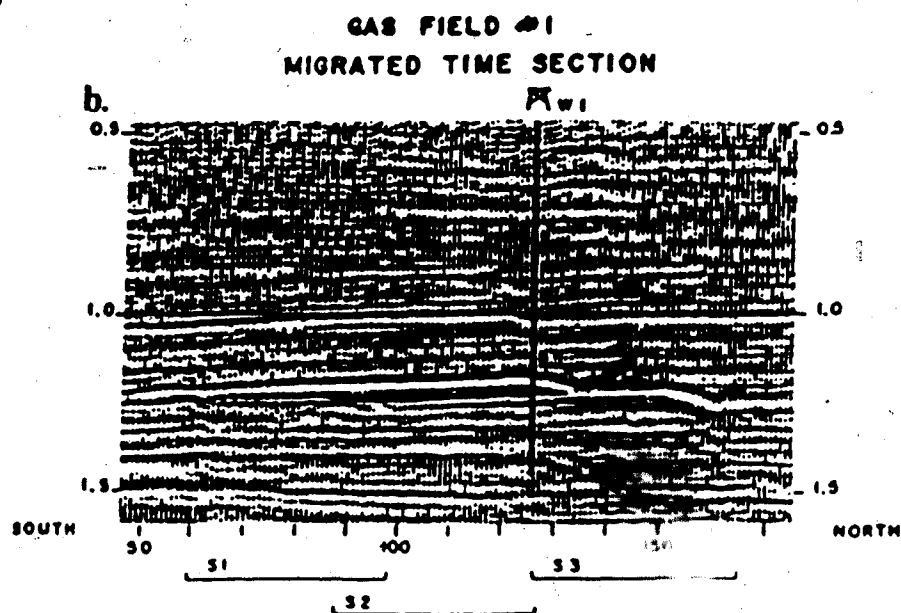
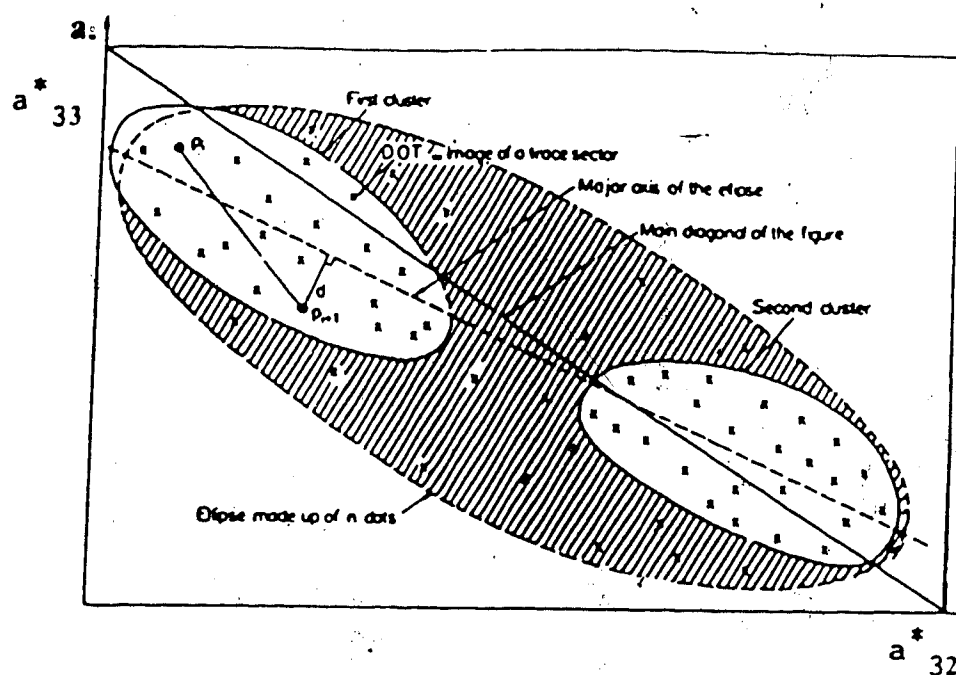


Figure 14.... a. Cartesian cluster graph, b. migrated seismic section (modified from Bois, 1980).

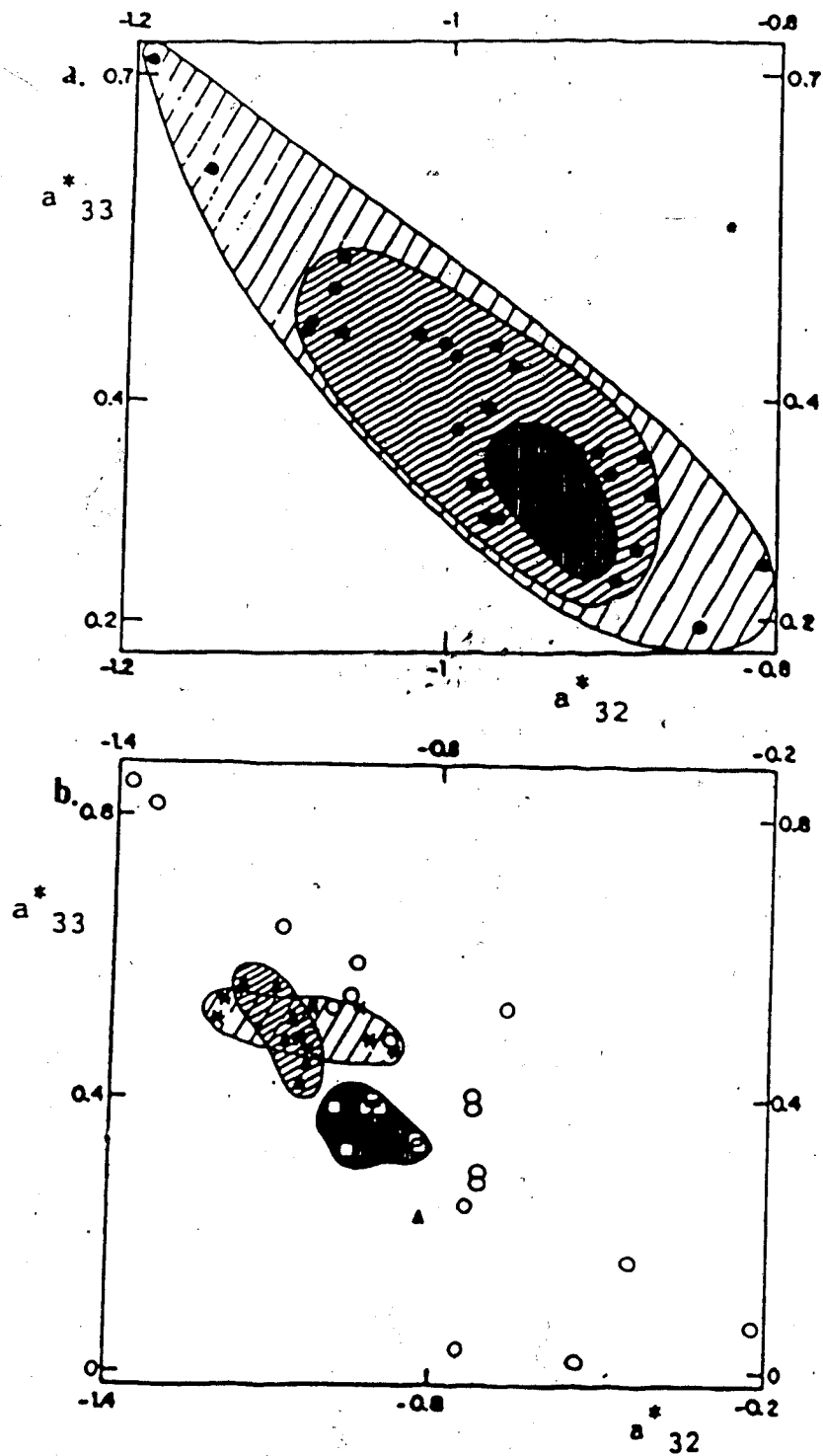


Figure 15.... a. cluster graph for sector S2 in b, b. cluster graph for sector S3 in b (modified from Bois, 1980).

generated. However, the approach still had difficulty as stated by Bois(1982). The problem was that a clear-cut clustering between the dots rarely occurred and some dots were poorly located as could be seen in Figure 15b.

3.2 Applications of Syntactic Pattern Recognition

One of the earliest attempts to use simple features such as peak location to do automatic seismic reflection picking was given in the paper of Paulson and Merdler (1968). The picking scheme had two main steps. The first step was to detect peaks within a trace. The second step was to correlate laterally the peaks with the peaks in the neighboring traces. Suppose P_{n,t_0} , the peak of the n^{th} trace at time t_0 , was tested for lateral continuity. In trace $n+1$, the time interval (t_0-d, t_0+d) was examined to detect the peak whose time position was closest to t_0 . The pre-determined parameter, d , defined the maximum dip within the section. Suppose P_{n+1,t_1} was the matched peak with the correlating peak P_{n,t_0} , then a straight line through the two peaks was extended to define a time t_2 in trace $n+2$. The time interval $(t_2-\Delta, t_2+\Delta)$ was again examined for matched peak, P_{n+2,t_2} . Δ was chosen to be smaller than d . A minimum of three peaks was necessary to define a reflection segment. Once three peaks were established, a least-square-fit straight line might be calculated and this line was extrapolated to next trace to find the matched peak. If such a peak existed, the procedure was repeated. If a peak could

not be found, the search was terminated and a new peak was chosen for a lateral continuity test. Output from this step could be displayed as a section of reflection segments. However, in order to pick out the significant segments for further analysis, each segment was assigned a grade which was the sum of the peak amplitudes in that segment. Then a threshold value was used to discriminate segments.

The scheme was applied to a seismic section of a salt dome which had been processed to give a high signal-to-noise ratio. The section showed good coherence of the flanks of the dome, but after the detection scheme, there was some loss of coherence because a series of short segments showed at those places. The scheme ignored local peaks within the negative parts of the trace and some of the termination of the lateral continuity search was certainly due to this factor. The other was the matched peak picking criterion which was based only on location. An inherent assumption in the scheme was that the best location was along the best fit straight line through three consecutive points. The correlating peak was matched with the peak which had the closest time position to this line. A problem of mis-match might arise when two peaks were on opposite sides of and close to the correlating peak. One was correct, while the other was not. However, since the correct one was a little bit farther away, it was excluded due to the matching criterion.

Liu and Fu (1982a & b) applied the syntactic technique to discriminate earthquakes from nuclear explosions. Each time series was segmented into N equal parts. The selection of the parameter, N was case dependent and was chosen as a compromise between representation accuracy and computational efficiency. Two features, namely zero crossing count, x_1 , and log energy, x_2 were extracted from each segment. Zero crossing count roughly represented the major frequency component of the signal in that segment; log energy indicated the magnitude of the signal. Each segment was then represented by a vector $X=(x_1, x_2)$ in the orthogonal feature space and classified into one of the clusters. Within each cluster, the distance between vectors and the vector of the cluster center was less than a threshold distance. The total number of clusters corresponded to the total number of primitives used in representing the signals. After all the signals were reduced to strings of primitives, classification used the string-to-string matching method.

String-to-string matching involved cost calculation. Since all the signals had the same number of segments, only substitution cost was calculated for each string. The test string was matched against the reference strings and then assigned to where the closest match occurred using the nearest-neighbour decision rule.

Liu and Fu (1982a & b) used seismic data recorded at LASA in Montana. Among them, 111 records were nuclear explosions and 210 records were earthquakes. 41 earthquake

and 59 explosion records were selected as the training or reference samples. Each record contained 1200 points. The sampling frequency was 10 points per second. Each record was divided into 20 equal segments with 60 points in each segment and 13 optimal clusters were used as primitives. Weights of substitution were the normalized distances between corresponding clusters. The analysis was done in VAX11/780 using Pascal language and 201 out of 221 events were correctly classified with an average time of 7 second for each record.

In this example, the assignment of weights to transformations was very important to the classification results. A constraint to the problem was that the records were segmented into the same number of portions to simplify the cost computation. It was suggested by the authors that the insertion or deletion transformations could be included by using the distances of the cluster centers from the origin as weights. However, the insertion and deletion transformations were not covered due to the complexity of their inclusion.

In 1982, Lu published a study of picking seismic reflectors using the string-to-string matching approach. Lu greatly reduced the data by ignoring the negative portion of the waveforms. In order to represent a trace as a string of peaks, each trace was scanned to detect peaks. During the detection process, a trace was decomposed into half "cycles" separated by zeros. A half "cycle" might contain one or more

peaks. A peak was regarded as an upslope followed by a downslope. In this way, peaks were detected and their features such as the location of the peak, LP, the duration of the peak, DU, the amplitude of the peak, AP, the average amplitude, AV, were extracted.

The detection procedure correlated the series of peaks on a trace with those on neighboring trace by comparing their feature set. The similarity between peaks a and b was measured by the cost or distance function

$$C(a,b) = W_1 * |LP_a - LP_b| + W_2 * |DU_a - DU_b| + W_3 * \log |AP_a - AP_b| + \\ + W_4 * \log |AV_a - AV_b| + (MDU - OVL) \quad (3.11)$$

where $W_1, W_2, W_3, W_4 = 0.5$ were weights. More discussion of each individual term will be given in section 4.2. The string-to-string matching algorithm determined the best match between the two traces as measured by the minimum cost of the sequence of transformations.

The scheme was applied to a stacked section with the result shown in Figure 16. Although the places of mis-matches were not mentioned, the result was appealing and suggested using syntactic pattern recognition techniques to automate the picking of seismic events. However, this example had a similar drawback to that in Paulson and Merdler's paper. The reduction of the redundancy of the data set by ignoring the negative portions of the trace was not

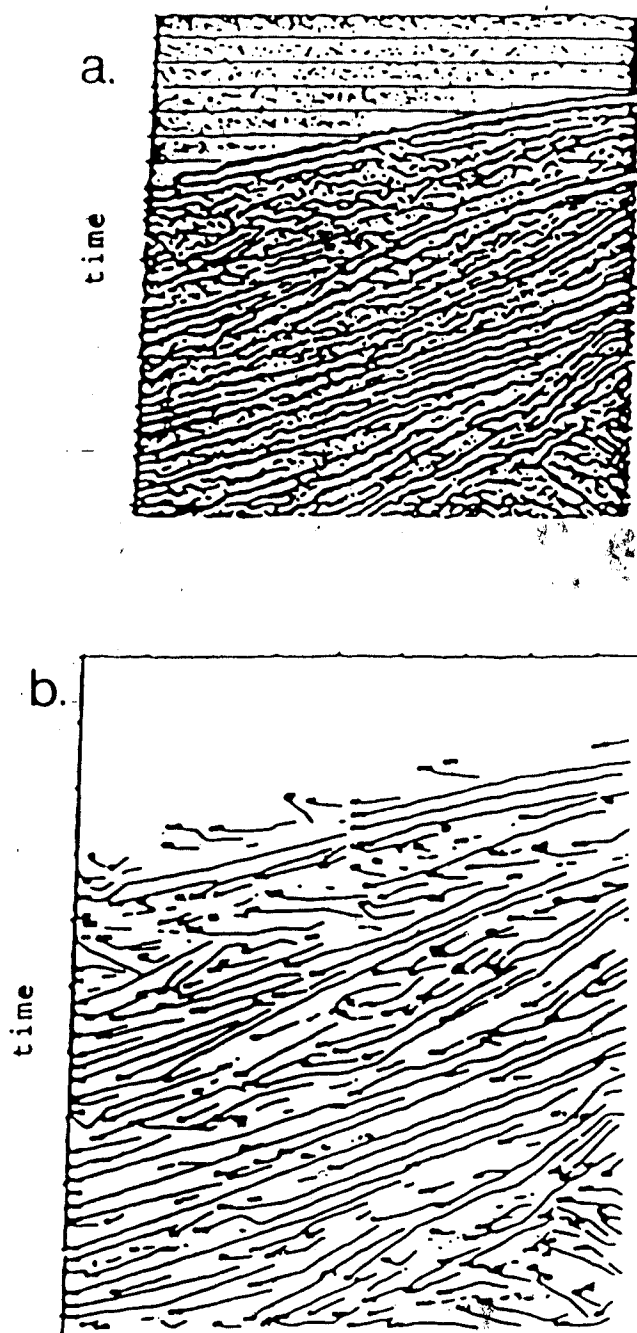


Figure 16.... a. A stacked section and b. result of pattern recognition (modified from Lu, 1982).

justified. Geophysically speaking, a trough was as important as a peak. Moreover, the logarithmic terms used in the cost function might possibly contribute to the total cost a negative partial cost which was harmful to the scheme. Also, the assignment of weights was not discussed.

From the analysis of the zero-phase Ricker wavelet, Huang and Fu (1983, 1984) proposed a tree classification method to detect bright spots in seismic signals. The major physical indicators associated with bright spots are high amplitude due to high reflection coefficients, low frequency due to high frequency attenuation, and polarity reversal due to a negative reflection coefficient at the gas-sand interface. Therefore, envelope, instantaneous frequency and polarity were used as features. Three kinds of hypotheses of bright spots were presented.

- a. The bright spot has high amplitude, low frequency and polarity reversal.
- b. The bright spot has high amplitude and low frequency content.
- c. The bright spot has high amplitude and polarity reversal.

In the scheme, bright spots were detected if one of the three hypotheses was satisfied. The bright spot satisfying the first hypothesis was the most restrictive one.

For each input seismogram, some testing traces were selected to equally cover the reflection characteristics of every kind of layers. From these traces, a threshold of

envelope was first used to discriminate signal from noise. Instantaneous frequency could be applied to separate a low frequency zone from a normal zone. The threshold of instantaneous frequency was determined by an inspection of good separability of the data in the scattering diagram. A scattering diagram was a display of the signal's envelope versus instantaneous frequency. If possible, polarity could be used instead of instantaneous frequency, to detect the reflection in a gas and oil sand zone.

Real data was shown in Figure 17a in which the dominant wavelets were the zero-phase wavelets. From the scattering diagram (see Figure 17b), the low frequency was found to be not significant. A classifier of the third hypothesis was designed. The classification result was shown in Figure 17c. The bright spot at 1.7 sec most probably contained gas. At the middle part of the layer of 1.7 sec, the reflection of positive polarity indicated the probability of oil/water. Without using the classification technique, a bright spot might be detected at the middle part of the layer at 1.7 sec by visual inspection.

Gaby and Anderson (1984) employed affinity to analyze the onset of P- and S- waves as part of a syntactic pattern recognition system to automate the interpretation of seismic signals. Affinity is a hierarchical clustering technique which builds a tree-like description of the waveform, thus providing a powerful tool for investigating the morphological structure of the data.

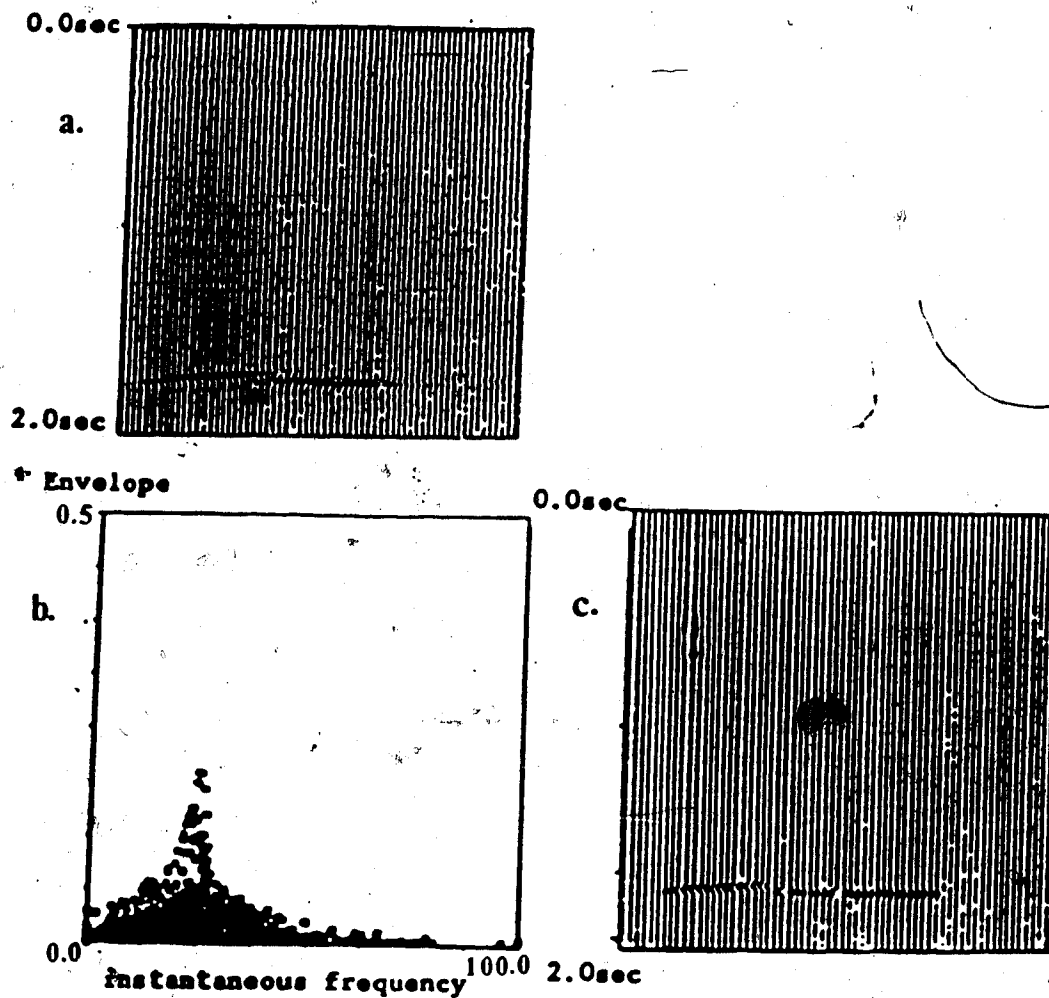


Figure 17.... a. Relative amplitude seismogram of bright spots. b. Scattering diagram. c. Classification result (modified from Huang and Fu, 1984).

Affinity requires three main components. The first is the initial segmentation of the seismic signal with each segment described by a feature vector. Second, a linking function links similar adjacent segments. A similarity or distance function, F , compares the feature vector $x(i)$ with its neighbors $x(i-1)$ and $x(i+1)$. If $F[x(i), x(i-1)] < F[x(i), x(i+1)]$, then segment i is linked to segment $i-1$; otherwise, it is linked to segment $i+1$. The last is the merging function which merges the two linked segments to form a higher node. In this study, the merging criterion requires that two segments A and B be merged if and only if segment A is linked to B and segment B is linked to A . This is the so-called "double links" criterion.

Amplitude and slope are used as features since a rapid increase of these features is usually associated with the onset of P- and S- wave.

A. Average Amplitude Criterion

Each segment is represented by a feature vector

$$x(i) = [L(i), A(i)] \quad (3.12)$$

where $L(i)$ is the length and $A(i)$ is the average amplitude of the i^{th} segment. The linking function for $x(i)$ and $x(i+1)$ is

$$F[x(i), x(i+1)] = \frac{A(i)L(i) - A(i+1)L(i+1)}{L(i) + L(i+1)} \quad (3.13)$$

The merging of these segments produces a new segment M whose length is

$$L(\text{new segment}) = L(i) + L(i+1) \quad (3.14)$$

and average amplitude is

$$A(\text{new segment}) = \frac{A(i)L(i) + A(i+1)L(i+1)}{L(i) + L(i+1)} \quad (3.15)$$

B. Slope Criterion

Each segment is represented by a feature vector

$$x(i) = \{[T(i), A(i)], [T(i+1), A(i+1)]\} \quad (3.16)$$

where $A(i)$ and $A(i+1)$ are the average amplitude of the i^{th} and the $(i+1)^{\text{th}}$ segments and $T(i)$ and $T(i+1)$ are the start and end point of the i^{th} segment. The criterion compares the two slopes, and links to adjacent segment with the smallest change in slope. The linking function for $x(i)$ and $x(i+1)$ is

$$\begin{aligned} F[x(i), x(i+1)] = & A(i) \frac{T(i+1) - T(i+2)}{T(i+2) - T(i)} \\ & + A(i+1) \\ & + A(i+2) \frac{T(i) - T(i+1)}{T(i+2) - T(i)}. \end{aligned} \quad (3.17)$$

The merging of these segments produces a new segment M , where

$$M[x(i), x(i+1)] = x(\text{new segment})$$

$$= \{[T(i), A(i)], [T(i+2), A(i+2)]\}. \quad (3.18)$$

A depth 16 tree structure of one waveform derived from the average amplitude criterion was shown in Figure 18a. Regions with high similarity were correctly merged to form larger segments. The generation of morphologies from a tree structure could be defined by (1) choosing a maximum depth in the tree, and (2) choosing a maximum segment length.

The first method of morphology extraction was shown in Figure 19A. In each, the original signal was shown in dashed lines and the resulting signal was in bold lines. At high levels of the tree, the resulting morphology missed the P-wave onset. As the maximum depth was increased, more detail was revealed. The second method was demonstrated in Figure 19B. The sequence of structure was determined by tracing down the tree until the segment length was less than a threshold value or the bottom of the tree was revealed.

A depth 16 tree structure of one waveform derived from the slope criterion was shown in Figure 19C. The first extraction method was shown in Figure 20A. At

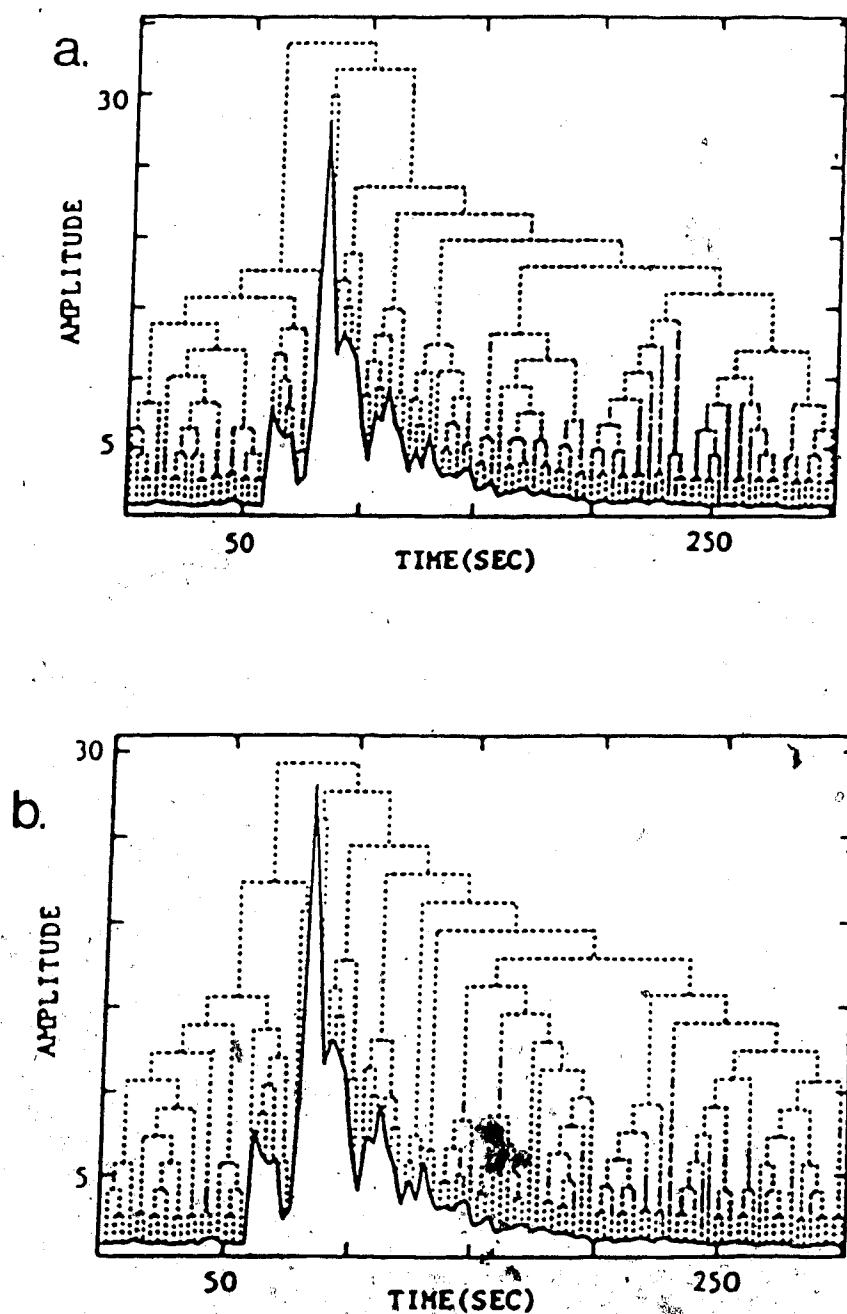


Figure 18.... Tree structure generated by affinity algorithm based on a. average amplitude criterion and b. average slope criterion (Gaby and Anderson, 1984).

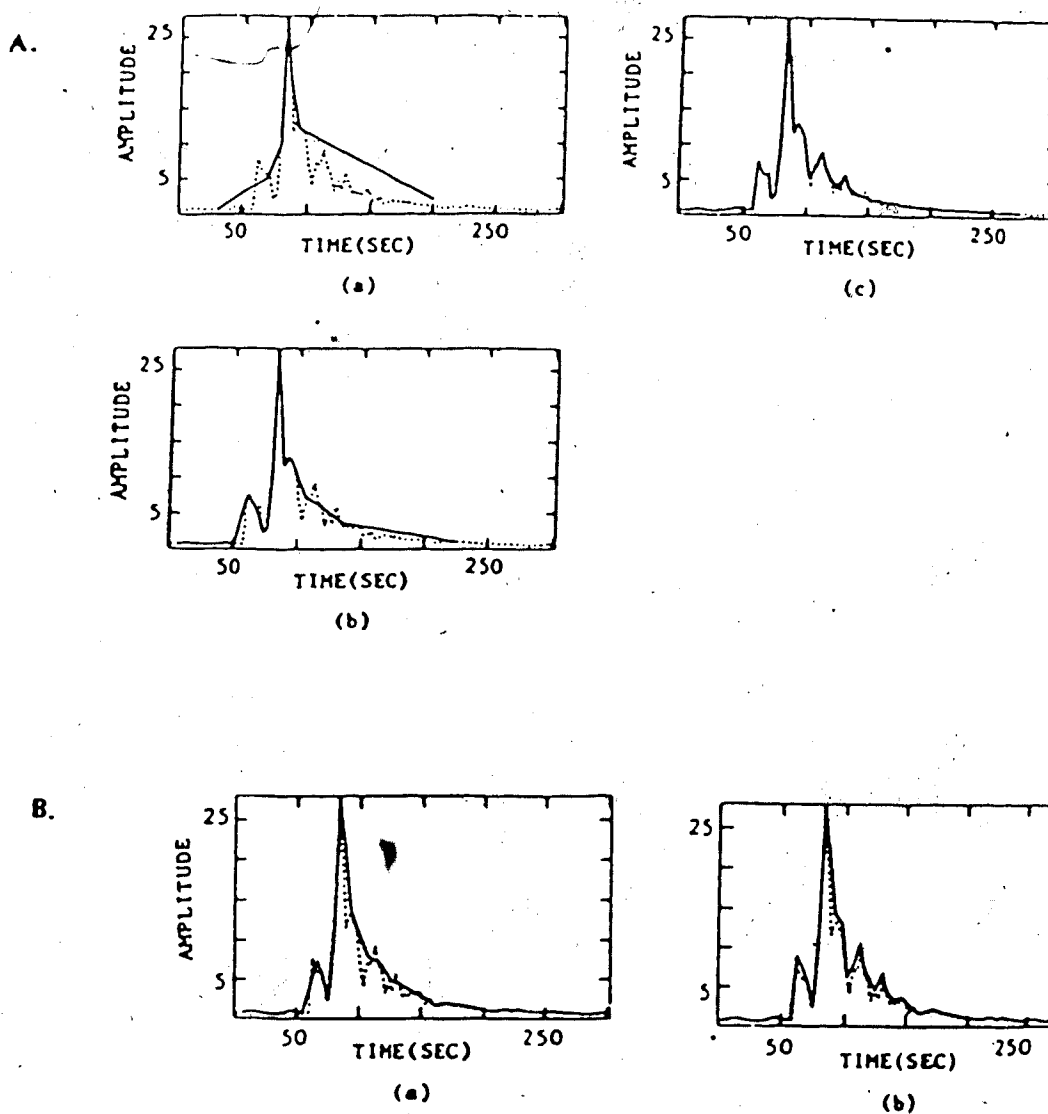


Figure 19.... Resulting morphologies based on the average amplitude criterion. A. Results obtained from maximum depth: a. depth of 3; b. depth of 5; c. depth of 6. B. Results obtained from maximum segment lengths: a. length of 18 sec; b. length of 12 sec; c. length of 8 sec.

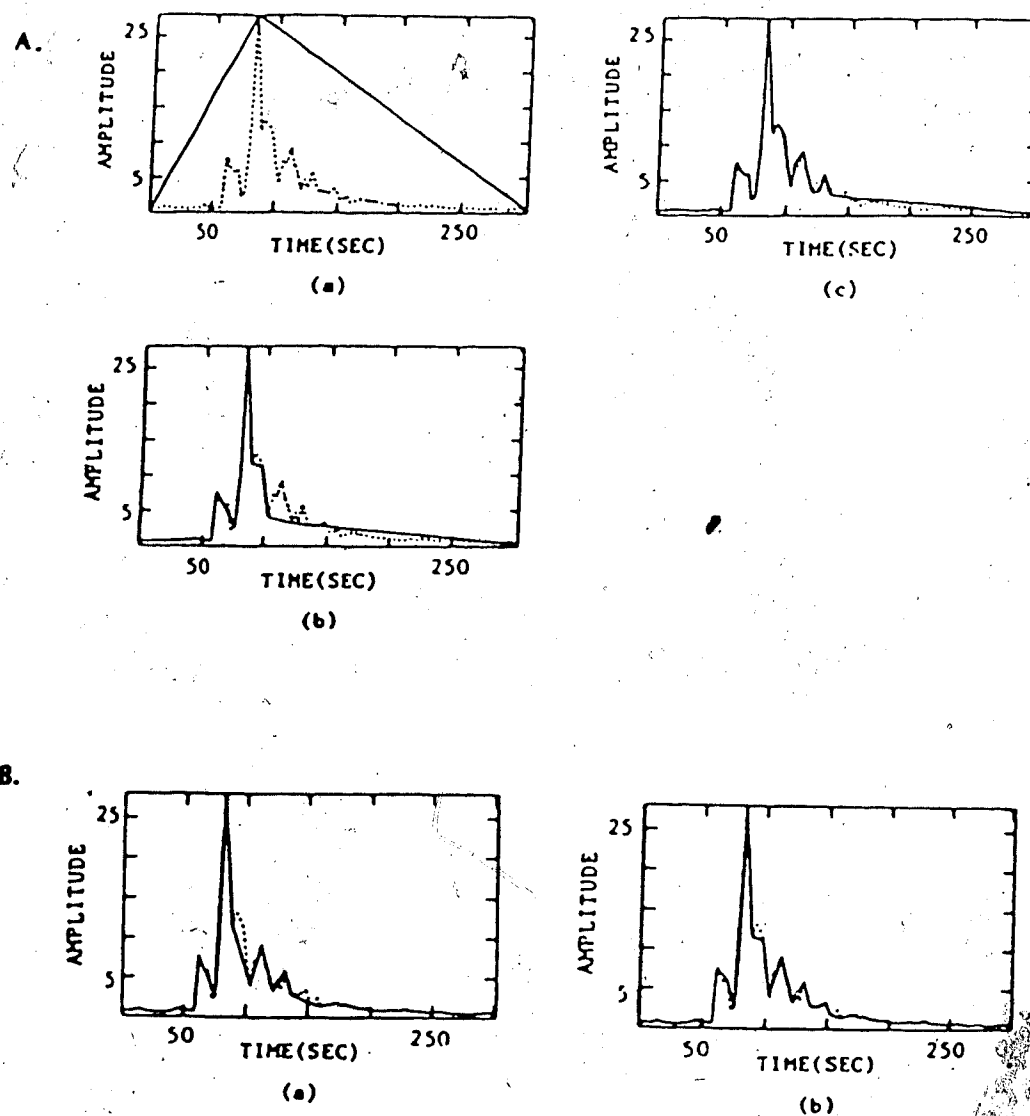


Figure 20.... Resulting morphologies based on the slope criterion. A. Results obtained from maximum depth: a. depth of 1; b. depth of 4; c. depth of 6. B. Results obtained from maximum segment lengths: a. length of 18 sec; b. length of

first levels of the tree, the resulting morphology picked out the S-wave but missed the P-wave onset. As the maximum depth was increased, more detail was revealed. The second method was demonstrated in Figure 20B. The resulting morphologies picked out the true location of the P- and S-wave onsets.

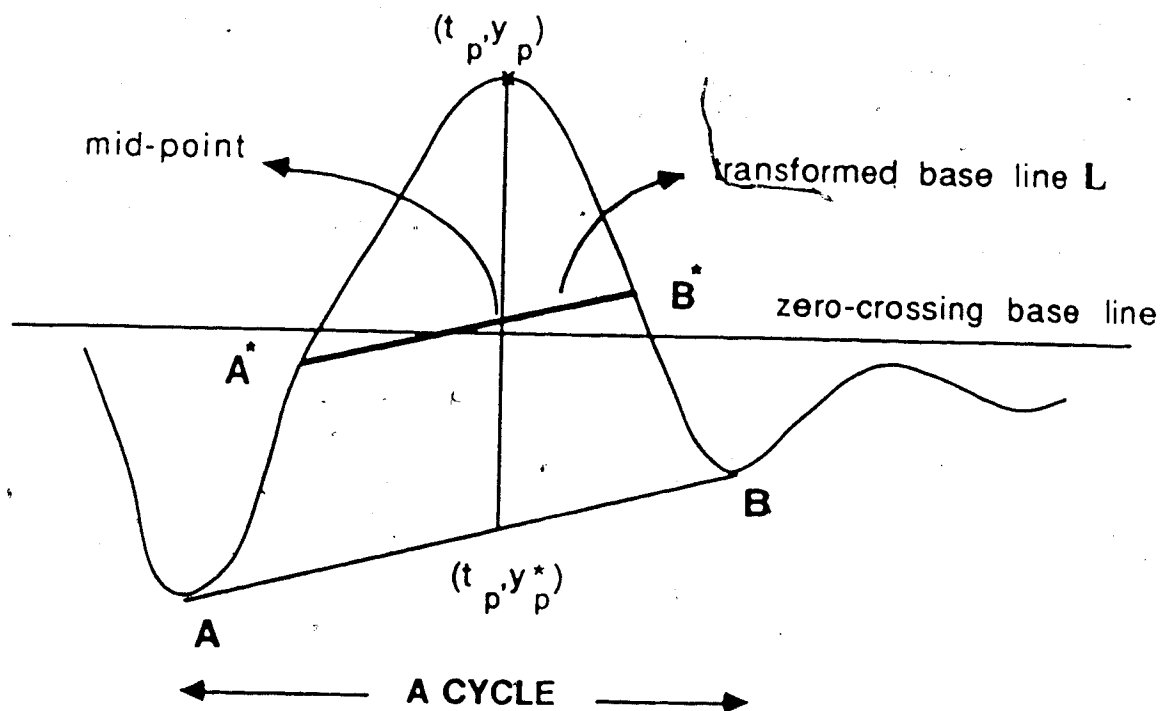
4. An Automated Seismic Event Picking Procedure

4.1 Primitives and features extraction

Primitive extraction is the first step in syntactic pattern recognition. Let M be an $S \times T$ matrix in which each element M_{st} is an amplitude of trace s at time t . This matrix is input for primitive detection. In the procedure, I consider each trace as a series of cycles (see Figure 21). Each cycle is a waveform bounded by two consecutive local minima. Peaks of cycles are primitives in the experiment. A peak is defined as an upslope followed by a downslope and its occurrence time t_p , is a feature and is recorded. Unlike Lu's scheme (1982) which only used the positive peaks, I consider all local peaks as candidates for correlation.

In order to extract features of all the primitive peaks, including the local peaks within the negative portion of a trace, the original zero-crossing base line has to be transformed. In Figure 21, each data point of the waveform is represented by coordinates (t, y) where t denotes time and y , amplitude. AB is a segment linking two local minima, each point along AB is represented by (t, y^*) , where y^* values are interpolated linearly. The transformed peak amplitude, $\hat{y}(t_p)$ is defined as

$$\hat{y}(t_p) = \frac{y(t_p) - y^*(t_p)}{2} \quad (4.1)$$



$$\hat{y}(t_p) = \frac{y(t_p) - y^*(t_p)}{2}$$

Figure 21.... A cycle within a trace.

The transformed zero-crossing line A^*B^* , is parallel to AB at $[t_p, y(t_p) - \hat{y}(t_p)]$ where A^* and B^* are the intercepts of L with the waveform. The duration of the peak is the time difference between A^* and B^* . If each waveform is sinusoidal, then A^*B^* overlaps the original zero-crossing line and our definition is consistent with the classical one. Output from this part is a $3 \times N \times S$ array P containing all the peaks of S traces with their three features. The elements $P_{1,n,s}$, $P_{2,n,s}$ and $P_{3,n,s}$ denote amplitude, $\hat{y}_{n,s}$, arrival time, $t_{n,s}$ and duration, $D_{n,s}$ of the n^{th} peak of the s^{th} trace.

4.2 Record analysis

Seismic record analysis involves recognition of the major coherent events, extraction of these events' features and the analysis of record structure in terms of these events. Following the logic of human visual interpretation, the procedure can be divided into three steps. The first step deals with trace-to-trace correlation in which peaks of one trace are matched with those in the neighboring trace. In the next step, the matching pairs are connected to form sets of events. Coherent events are picked from this step and assigned confidence values. The third step is zonation. Zoning is a partition of record and divides a record into groups or zones of coherences.

4.2.1 Trace-to-Trace Correlation

In a geological environment where the stratigraphy is extremely simple and all interfaces are parallel, cross-correlation is a good technique for correlation between seismic traces. It has general applicability and requires no more subjective judgements than the decision on the window length. In normal geology, the technique fails since it does not account for gaps, missing intervals, or abrupt thickness variations of strata. Modifications have been suggested to improve the correlation quality in such situations, but problems still exist, since the nature of the variation is sometimes unidentifiable.

Shaw and Cubitt (1979) proposed stretching the shorter zones of strata to account for the variation of thickness due to differing rates of sedimentation. This stretching is done before cross correlation and after each log has been segmented into zones. Cross-correlation without stretching provides matches in the presence of faulting. The larger value of the two correlation results indicates the nature of the variation. Judgement made from these values is usually not valid, because the correlation values also depend upon the number of data points which are in phase within the window. Therefore, Robinson (1978) commented that, the approach could lead to errors when points of correlation were unknown.

The choice of window length is also an inherent difficulty in the design of the cross-correlation scheme.

The number of possible correlations is a function of the window length. Therefore, the position of the best correlation changes with the length and position of the correlation window. The dominant peak in the window also affects the correlation result. Indeed, cross-correlation is unable to resolve short intervals consistently because the statistical significance of the correlation coefficients depends on the number of independent observations. Another problem with the scheme is the possibility of crossing of correlation events. Such a result is peculiar since lithologic sequence seldom reverses over short distance.

In a knowledge-based system, the cross-correlation technique which only uses the statistical measures of the data without using any a priori information in the analysis must give way to string-to-string matching. The latter approach is also called dynamic waveform matching and allows distortion of two matching sequences due to insertion, deletion and substitution transformations. The scheme uses the whole sequence as input, thus eliminating the difficulty with window-length. The matching is made sequentially. Since the cost is a measure of dissimilarity between two sequences and is derived at any point from the previous total cost, the correlation events never cross. Wu and Nyland (1986) used the string-to-string matching algorithm to do well log interpretation and described the technique as a framework for automated borehole to borehole correlation and stratigraphic interpretation, because much geological and

geophysical knowledge of the area under study and the data relevant to the borehole being considered could be incorporated into the cost function.

In my analysis procedure, the array P from the primary step is used for correlation. Any two adjacent traces are correlated by string-to-string matching. Four types of constraints discussed by Anderson and Gaby (1983) are imposed on the algorithm.

A. Endpoint constraint

For any two matching sequences, the transformation path begins at $i=j=0$ where null primitives exist. The last matching point is within 0.1 sec after the end of the shorter sequence.

B. Local continuity constraint

For each grid point (i, j) in the matching grid, only three possible paths (see Figure 5) are used to derive its associated distance $D(i, j)$ from those of its neighbors. This constraint ensures that the optimal path is nondecreasing and has the least cost.

C. Global path constraint

A matching peak to a peak of one trace is sought within a 0.2 sec window of the other trace for a matching peak. The window is centered at the location of the peak in the first trace. The limitation of the search range eliminates some of the mis-matching errors and reduces computing time.

The constraint is also justified here in that faulting is not expected in the area under study.

D. Local cost measure

Anderson and Gaby (1983) required the distance calculated to satisfy the classical distance definition:

- 1) $d(x,y) \geq 0$,
- 2) $d(x,y) = 0$ if $x=y$,
- 3) $d(x,y) = d(y,x)$ and
- 4) $d(x,y) + d(y,z) \geq d(x,z)$

where x , y and z were three sequences. Since the distance is derived from the summation of individual transformation cost, I define the local cost measure:

- 5) $c(i,j;T_1) \geq 0$,
- 6) $c(i,i;T_S) = c(j,j;T_S) = 0$,
- 7) $c(i,j;T_i) = c(j,i;T_d)$ and vice versa, and
- 8) $c(i-1,j;T_i) + c(i,j;T_d) \geq c(i,j;T_S)$
or $c(i,j-1;T_d) + c(i,j;T_i) \geq c(i,j;T_S)$

where $T_1 \in \{T_i, T_d, T_S\}$ and i and j denote any two matching primitives. The requirement D.8 is most difficult to satisfy in this algorithm. The fulfillment of this criterion requires a careful

design of the cost function and assignment of weights. Nevertheless, the result of matching is not very sensitive to this error. Mismatches due to this factor can be amended by subsequent refining analysis.

The crucial part of the string-to-string matching algorithm lies in the evaluation of the cost function. All features that are useful in the comparison are considered in the function. The cost function embodies the knowledge about the physical characteristics of the objects being compared, in this case, waveforms and an interpreter's criteria of the event correlation when the correlation is done manually. In this particular example, the cost function used is

$$\begin{aligned}
 c(i, m; j, m+1) = & W_1 * |t_{i, m} - t_{j, m+1}| \\
 & + W_2 * |D_{i, m} - D_{j, m+1}| \\
 & + W_3 * \left| \frac{\hat{y}_{i, m} - \hat{y}_{j, m+1}}{(\hat{y}_{i, m} + \hat{y}_{j, m+1})/2} \right| \\
 & + W_4 * (\text{MDU-OVL})
 \end{aligned} \tag{4.2}$$

for substitution and

$$\begin{aligned}
 c(i, m; j, m+1) = & W_1 * |t_{i, m} - t_{j, m+1}| \\
 & + W_2 * |D_{i, m} - D_{j, m+1}|
 \end{aligned}$$

$$+W_3 * \left| \frac{\hat{y}_{i,m} - \hat{y}_{j,m+1}}{p_i} \right|$$

$$+W_4 * (\text{MDU} - \text{OVL}) \quad (4.3)$$

for insertion or deletion. In the functions, W_1 , W_2 , W_3 , and W_4 are weights and

$$p_i = \frac{\sum_{i=m}^{m+1} \sum_{j=1}^{l_i} \hat{y}_{j,i}}{\sum_{i=m}^{m+1} l_i} \quad (4.4)$$

where l_i is the total number of peaks of trace i .

In equations 4.2 and 4.3, the first term accounts for the location difference; the second term, the duration difference; the third term, the peak amplitude difference with respect to mean of the two correlating peaks in case of substitution or mean of all the peaks of the two correlating traces in case of insertion or deletion. The fourth term accounts for the non-overlapping part of the two corresponding positive parts of the transformed waveform (see Figure 22). $\text{MDU} = \min(D_{i,m}, D_{j,m+1})$, and OVL is the overlap between them. For insertion and deletion, a peak is matched with a null primitive. A null primitive has zero amplitude and nonzero location and duration terms. Thus one of the amplitudes in (4.3) is zero.

There is a reason for the simplicity in choosing these cost functions. Human interpreters, even experts, frequently

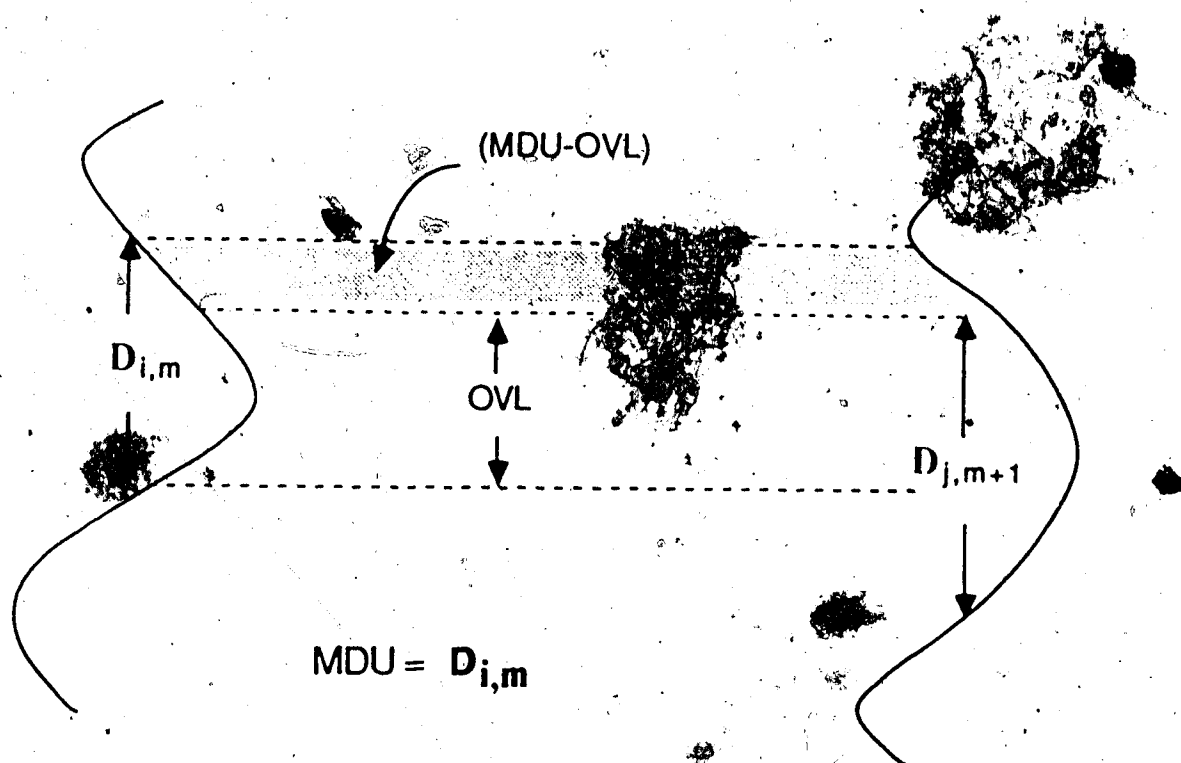


Figure 22.... The non-overlapping part of two waveforms.

compare two waveforms by noting their physical and temporal difference. The cost functions attempt to encode the human experience in comparison. Sophisticated logarithmic and exponential functions could be designed if more precise discrimination is needed, but the example illustrated in this thesis excludes this possibility.

Since the terms in the cost function are not normalized, the distribution of weights should be such that the average values of the terms indicate their importance. The mean term of Lu's cost function (see eq. 3.11) has been incorporated into my amplitude term by adding a constant $2/\pi$. The constant is obtained by averaging a sine wave given by $\sin\theta$ over a period of 2π . The weights are assigned to be 0.3, 0.2, $(1+2/\pi)$, and 0.1, respectively. The weights are chosen in such a way that the third term is emphasized.

Lu (1982) used $w_3 \cdot \log|(\hat{y}_{i,m} - \hat{y}_{j,m+1})|$ for the third term in both cost functions. This is harmful to the scheme since two peaks of very similar amplitudes will yield a negative infinite cost. The negative partial cost might cause the total distance of the optimal path to violate the local distance constraint D.1. Moreover, for the same difference in peak amplitude, pair of strong peaks should imply less cost than the weak pair does, this criterion is reflected in my cost function by 'normalizing' the amplitudes in the third term.

For each pair of correlated peaks in two consecutive traces in P, the string-to-string matching algorithm is

applied to compute the cost of all three types of transformations and to pick the correlation which has the least cost. A correlation array C of dimension $2XV \times M$ is defined. The elements C_{1vm} and C_{2vm} form the v^{th} matching pair of the m^{th} correlation in which the former is peak location from the m^{th} trace and the latter from the $(m+1)^{\text{th}}$ trace. Only substitution pairs are recorded in C . Deletion and insertion are considered discontinuities and are ignored.

It is beneficial, at this stage, to associate with each matching pair, a confidence of the correlation results. The confidence matrix F is defined as:

$$F_{vm}(i, m; j, m+1) = \frac{\sqrt{\hat{Y}_{i,m} * \hat{Y}_{j,m+1}}}{C(i, m; j, m+1)} \quad (4.5)$$

where F_{vm} is the correlation confidence of the v^{th} matching pair of the m^{th} correlation. The greater the peak amplitude and the smaller the cost of matching are, the greater is the confidence value.

4.2.2 Coherence analysis

The output of the trace-to-trace correlation is then used to detect the coherent events. An operation of tracing correlated peaks from trace to trace is run and all the correlated peaks are connected at their location to designate a coherent event, thus defining a $2XV \times Z$ array R in which elements R_{1yz} and R_{2yz} denote the trace number and the

corresponding peak location of the y^{th} component of the z^{th} coherent event. Useful information such as the extent and the average amplitude of each event are recorded as properties for further analysis.

A coherence confidence vector F of dimension Z is also attached to each coherent event in array R . Each element F_z is the sum of the correlation confidence of the correlation segments in C making up the z^{th} event:

$$F_z = \sum_{y=1}^Y P_{yz} \quad (4.6)$$

Both the correlation confidence and coherence confidence measure uncertainty about the decision made by the correlation algorithm. Coherence confidence reflects the reliability and importance of a coherent event. Coherent events consisting of correlation segments of high confidence are reliable. The importance of a coherent event lies in its continuity and its possession of high peak amplitude. Coherent events with high confidence, in our case, should be the major reflection events which are our main objects of investigation. Many coherent events result from such string-to-string matching. The introduction of the coherence confidence as a measure of importance discriminates the major events from the rest in the record. This eliminates some of the worries that human interpreters have about the autopicking result discussed in section 1.3. The coherent events are sorted according to their confidence level and

coherent events with high confidence are selected to form a set of elements for zonation.

It is insufficient at this stage to pick coherent events based on trace-to-trace correlation only for at this stage some mismatches of peaks are unavoidable. Refining the cost function might help to reduce the mismatches but it is impossible to eliminate them all. The problem comes from limiting our analysis to only two adjacent traces in the first step. The expert's visual correlation is superior to the machine trace correlation mostly because his visual correlation is based on the whole record instead of two neighboring traces. I introduce an operation to examine and refine the validity of trace correlation generated to this stage.


The operation is parallel analysis, based on the slope of other correlation segments within a window. Essentially, the analysis takes into consideration the information beyond the features of the two matching peaks. By taking advantage of the surrounding slope, the algorithm's threshold of correlation judgement is made wider. The assumption is also justified since any sudden change in geology or topography should have its effect reflected generally.

In the analysis, an overlapping window m traces wide and t seconds long is moved across the record to search for mismatches. Within the window, the weighted mean and the weighted standard deviation, STD of the slope of the correlation segments are calculated. The weighted parameters

are obtained by weighting the slope and STD with the corresponding correlation confidence, F_{vm} . In this manner, the average slope within the window is mostly determined by the correlation segments of high confidence. A correlation segment is chosen as a possible mismatch if

$$1. \quad |\text{slope} - \text{slope}_m| > B_1 * \text{STD}, \text{ and}$$

$$2. \quad F_{vm}(i, m; j, m+1) < T_1.$$

If these conditions are satisfied, I seek a new  pair. This requires that a peak in the old matching pair has to be matched with a new peak of the other trace within an bounded interval. The new matching pair is accepted if three more criteria are satisfied simultaneously:

$$3. \quad |\text{slope}_n - \text{slope}_m| < |\text{slope}_o - \text{slope}_m| \text{ and}$$

$$4. \quad C_n(i, m; j', m+1) < C_o(i, m; j, m+1) \text{ or}$$

$$C_n(i', m; j, m+1) < C_o(i, m; j, m+1) \text{ and}$$

$$5. \quad C_n * |\text{slope}_n - \text{slope}_m| < B_2 * C_o * |\text{slope}_o - \text{slope}_m|$$

where subscripts m , n , and o denote average, new and old value respectively and i' and j' , the new peak of the m^{th} and $(m+1)^{\text{th}}$ trace. B_1 and B_2 are predetermined constants. The criteria can be made stringent by choosing small values for the constants. The threshold, T_1 is chosen to be in the lowest correlation confidence level, because most of the mismatches have low correlation confidence. The values used

for T_1 , B_1 and B_2 are 0.45, 1.0 and 0.5 respectively. Crossing of correlated events is prohibited. A new match is also added if the slope of the pair of uncorrelating peaks satisfies the criterion:

$$6. \text{ slope} \leq \min\{(B_3 * \text{STD}), T_2\}$$

where B_3 and T_2 are 2.0 and 0.008 respectively. The value of T_2 is determined by considering that two peaks must be close together if they are to be joined.

4.2.3 Zoning

For further study and data reduction, a seismic record can be divided into zones. Zones are characterized by homogeneous features within zones and significant differences between zones. Coherence confidence is taken as the major feature of zones. An event of high confidence and a group of adjoining coherent events with high confidence are likely to be the seismic responses arising from similar strata. I take these as zones and the regions of a record between them are also taken as zones. In this manner, a seismic record is represented by a string of zones with features. All these zones might have stratigraphic significance and can be traced from record to record.

The partition of zones described below is a simple clustering procedure. All coherent events with highest confidence level are collected into a set to form cluster

centers. Zoning is begun at the event of the highest confidence in the set. It is the 'seed' or clustering center of the first zone and then eliminated from the set. An event will join this zone if and only if the following criteria are satisfied:

- a. It is a member of the set, and
- b. It lies within T_3 seconds close to the zone, and
- c. No more than N coherent events with a total length less than L lie between it and the 'seed' of the zone.

Here T_3 is given a wider threshold which is 0.1. Each correlation segment is considered to have unit length. The maximum length of a continuous event is 48 units in length (there are 48 traces for each shot). therefore, N and L are chosen to be 4 and 50% of the maximum length of the zone. If the event meets the criteria, it is then deleted from the candidate set. The zone keeps growing until no coherent event can join the zone. Another coherent event of the highest confidence in the candidate set is chosen as the 'seed' of a new zone. If the set is not empty, the process repeats. If no other events satisfy the criteria, the seed is then considered as a zone by itself.

The region between two zones with high confidence coherence is also defined as a zone, in spite of the fact that it may consist of a collection of low coherent events.

Zone features such as average length, average confidence, average peak amplitude, average correlation confidence, and average complex attributes of the coherent segments within zones are extracted.

4.3 An illustrative example

Common shot gathers have been used as illustrative examples to clarify these ideas. In 1984, the Geophysics 528 class of the University of Alberta conducted a magnetic, gravity and seismic survey near the town of Allan, 50 km south-east of Saskatoon, Saskatchewan (see Figure 23). Two seismic lines (see Figure 24) were shot and the survey acquired approximately 6 miles of 400% coverage seismic data. The equipment used was a DFSV recording instrument with 48 channels using dynamite as source. Each trace was recorded for 2.04 sec with a sampling interval of 2 msec. Geophone 25 approximately corresponded to the uphole geophone. The parameters of the survey were 110 feet for geophone-group interval and 660 feet for shot interval. The data processing procedure involved a raw muting of the first part of each trace, variable gain equalization with a overlapping sliding window of 300 data points and band-pass filtering with an eight-pole Butterworth filter.

Three shot gathers shown in Figure 25, Figure 26 and Figure 27 are taken from the north-south seismic lines, B in Figure 24. The field records are studied thoroughly in an attempt to identify and follow certain reflection events

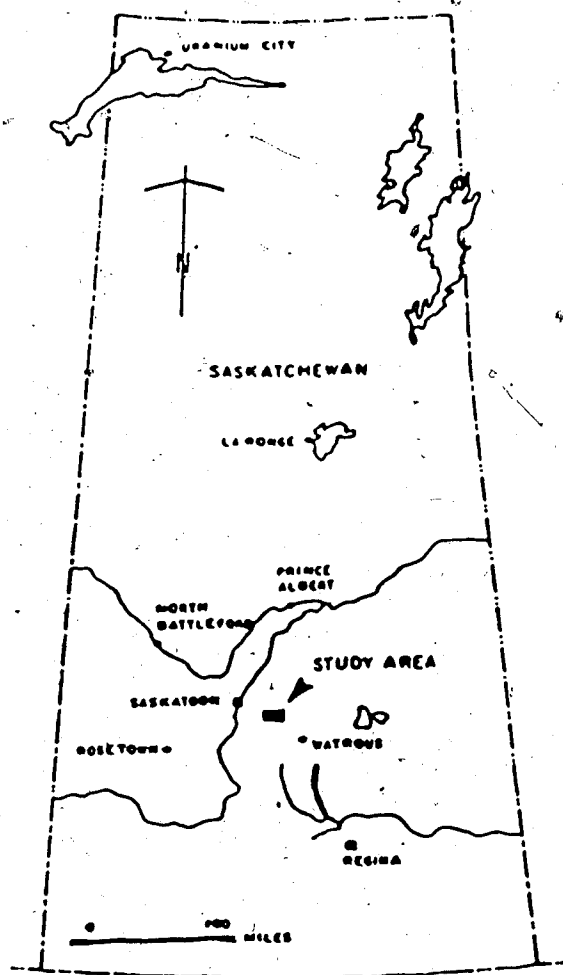
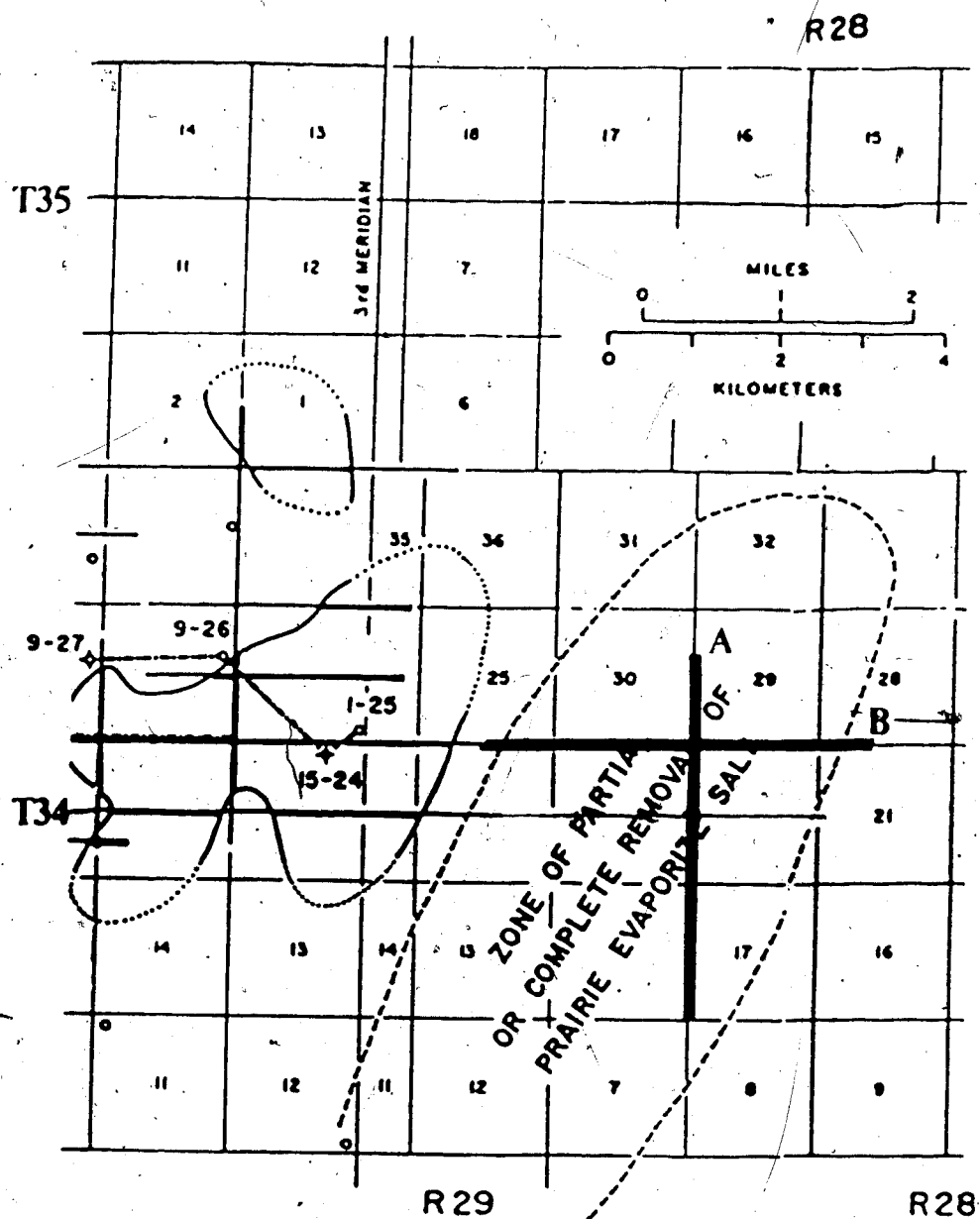


Figure 23.... Location map of the area under study (Gendzwill, 1978).



SEISMIC LINE ———
 REEF ANOMALY ———
 ASSUMED MOUND EDGE ———
 PROJECTED MOUND EDGE
 SECTION ———

DRILL HOLES :
 PARTIAL PENETRATION OF SALT °
 PENETRATED SALT +
 NO MOUND ◊
 FOUND MOUND +

Figure 24.... Plan map showing two seismic lines (A & B) surveyed by Geophysics class 528 (modified from Gendzwill, 1978).

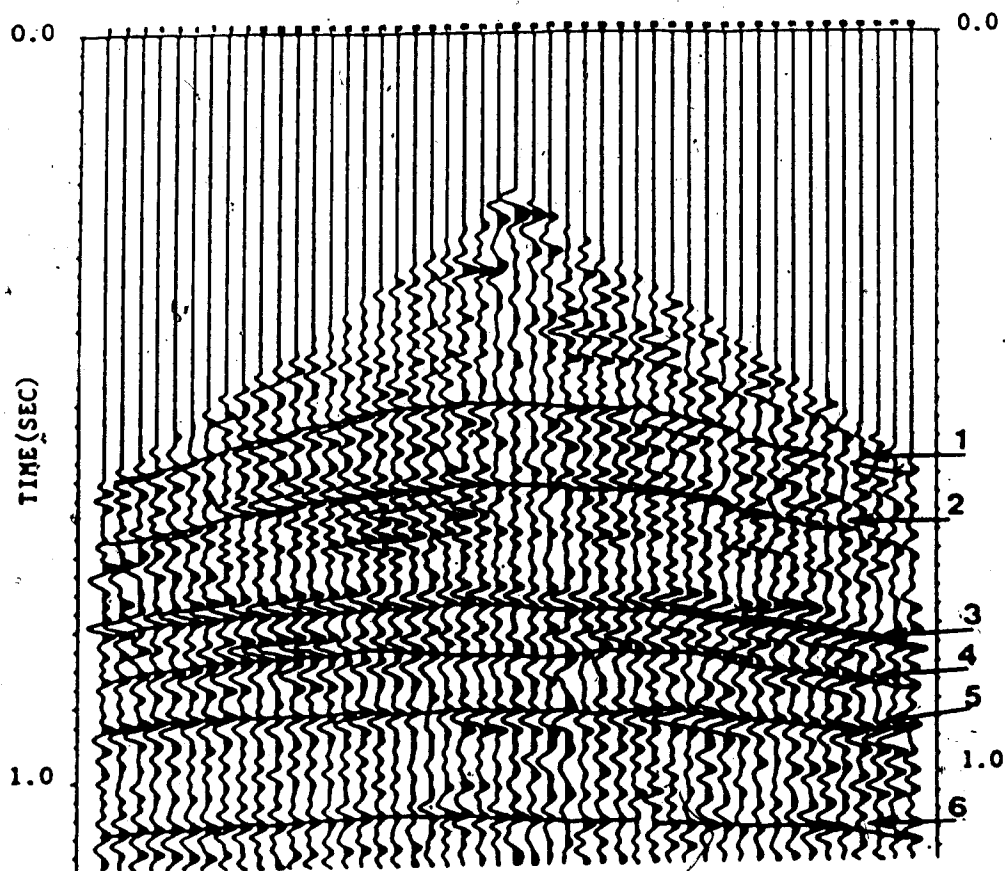


Figure 25... Seismic shot gather of shot 45 in which
1=Blairmore; 2=Souris River; 3=Prairie Evaporite;
4=Winnipegosis; 5=Cambrian; 6=Precambrian.

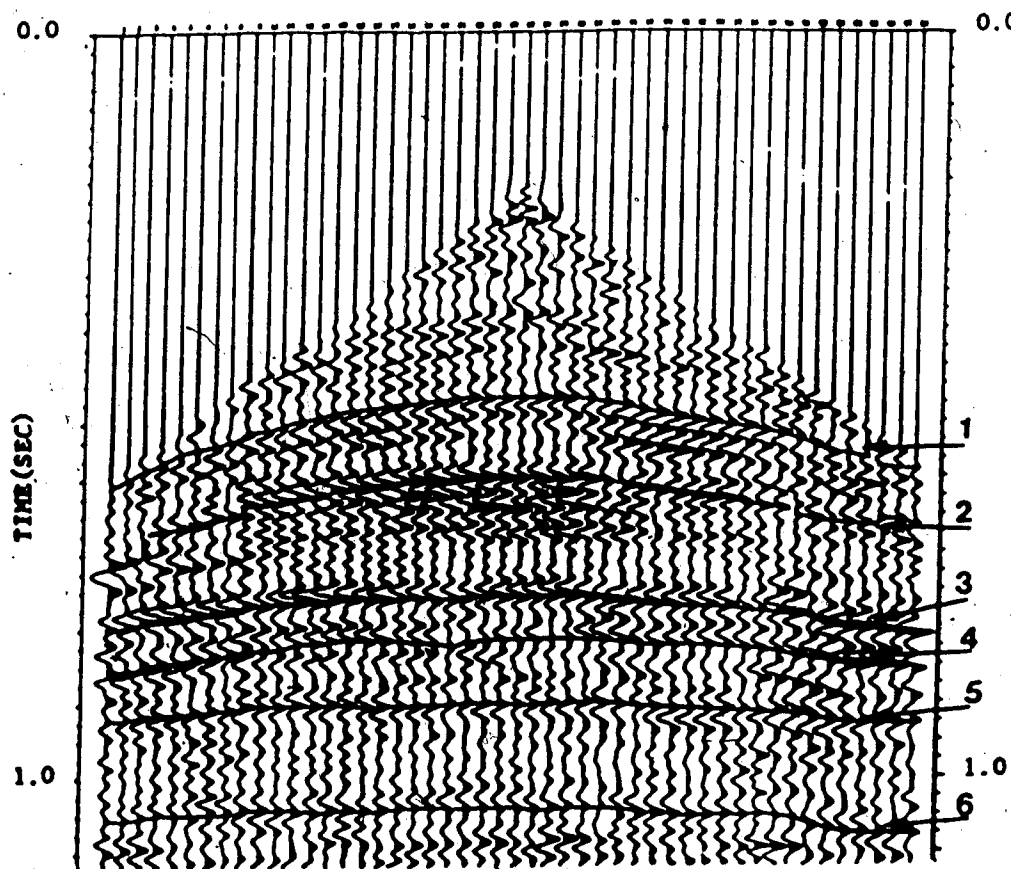


Figure 26.... Seismic shot gather of shot 46 in which
1=Blairmore; 2=Souris River; 3=Prairie Evaporite;
4=Winnipegosis; 5=Cambrian; 6=Precambrian.

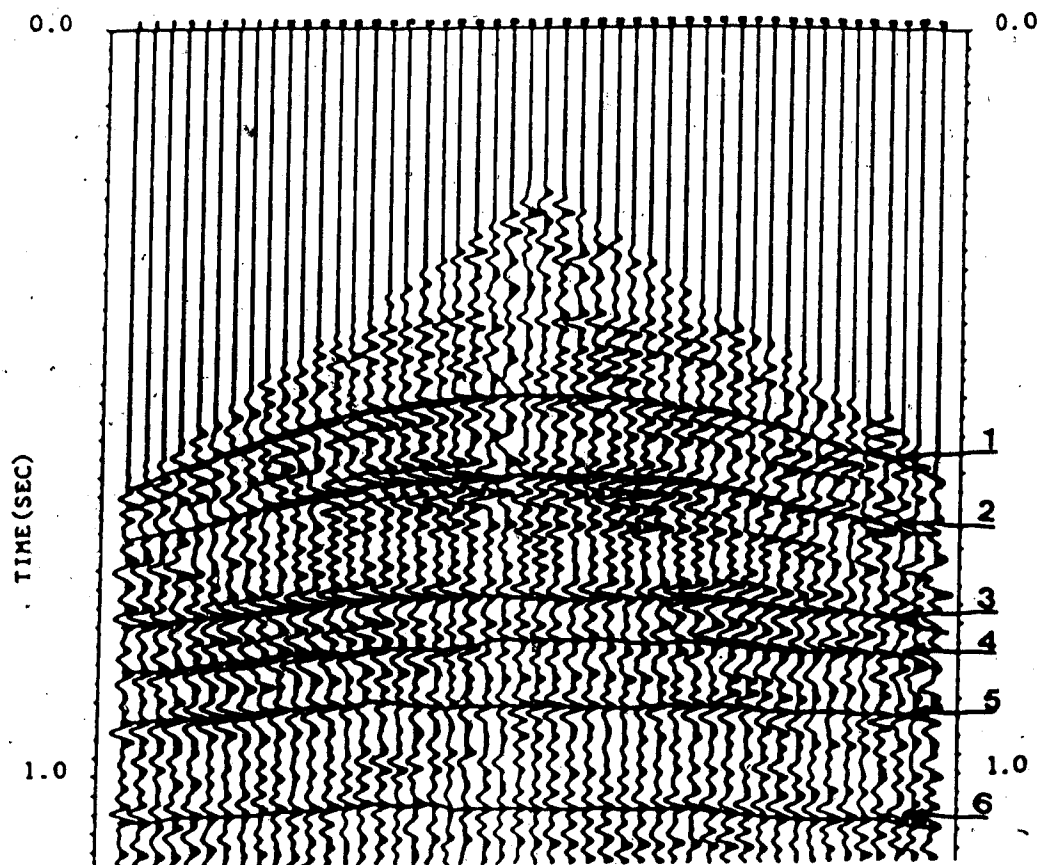


Figure 27..... Seismic shot gather of shot 47 in which
1=Blairmore; 2=Souris River; 3=Prairie Evaporite;
4=Winnipegosis; 5=Cambrian; 6=Precambrian.

from shot to shot. There are five fairly consistent reflection events (1, 2, 3, 4 and 5) which are identified to come from the approximate location of the tops of the Blairmore Formation, the Souris River Formation, the Prairie Evaporite Salts, the Winnipegosis Formation and the Cambrian. A sixth reflection event (6), which is not as consistent as the other, is speculated to represent the top of the Precambrian. The correlation of the events to the named formations has been crudely done through the use of the well logs and cross sections (Gendzwill, 1978). It is noticed that there is often a good event just before the Winnipegosis. This may correspond to the anhydrite bed or the Shell Lake Member referred to by Gendzwill. There are several cycles of high amplitude between the Prairie Evaporite event and the Winnipegosis event. Due to this factor, the recognition of Winnipegosis by string-to-string matching is difficult.

The matching procedure were applied to all 48 traces in the records. Figure 28 shows an example of the matching result of two adjacent traces represented by a series of peaks with their features. The thick line is the optimal correlation path. Figure 29 shows the trace-to-trace correlation result in record 45 as drawn in line segments (only the substitution pairs are shown). Most of the coherent events are detected. Part of the result using Lu's scheme is given in Figure 30 for comparison. My result shows fewer mismatches. Especially in places (A, B, C and D) where

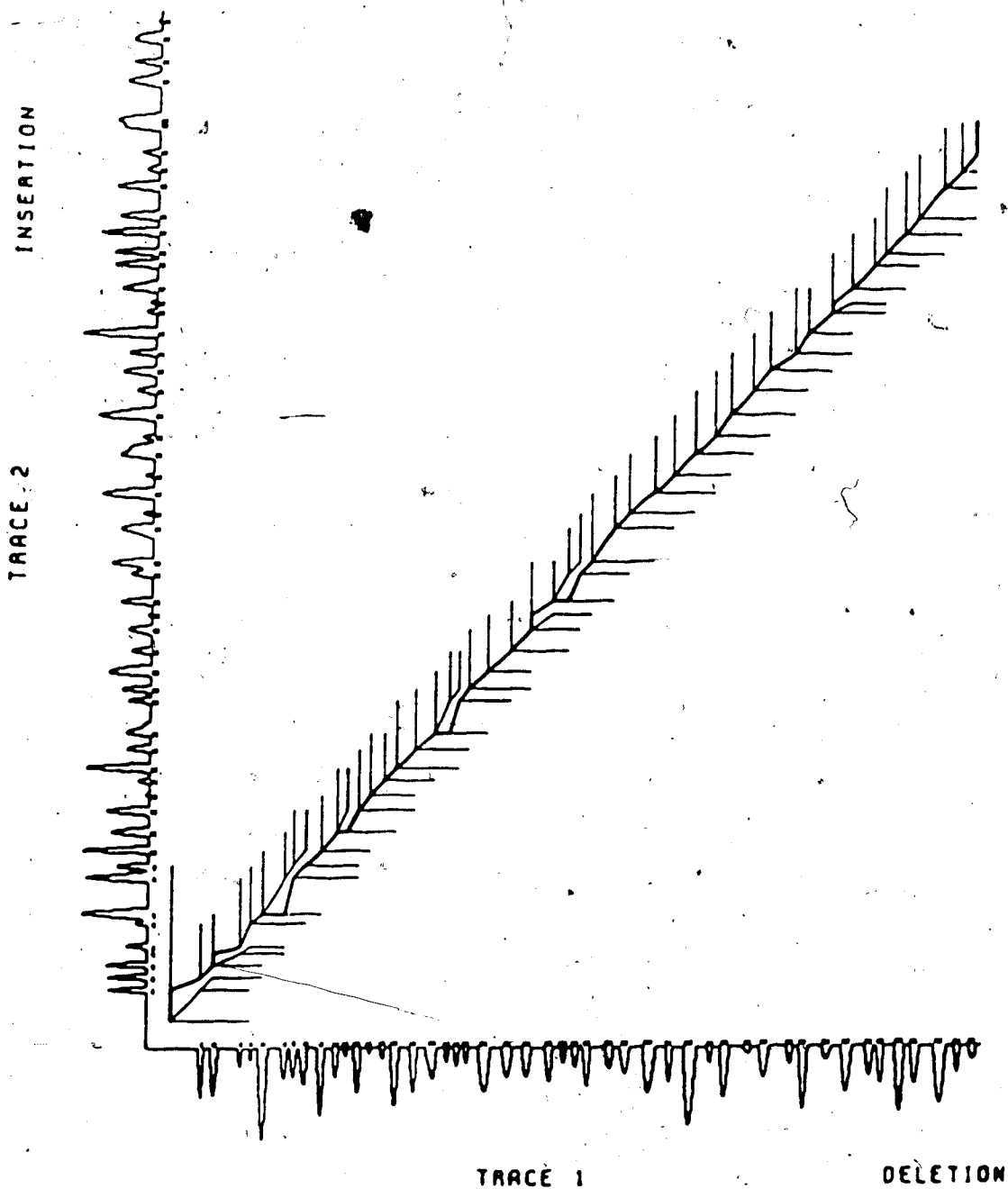


Figure 28.... Result of string-to-string matching of two neighboring traces.

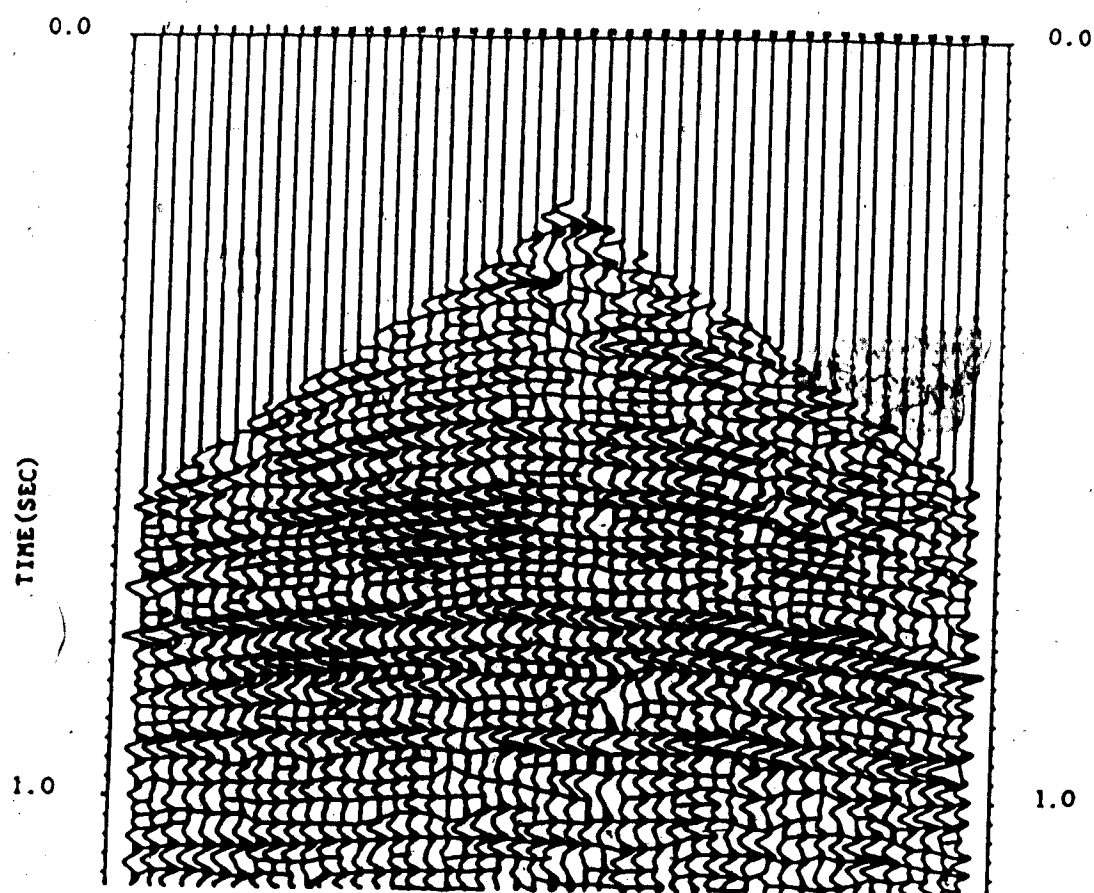


Figure 29.... Result of trace-to-trace correlation of record 45.

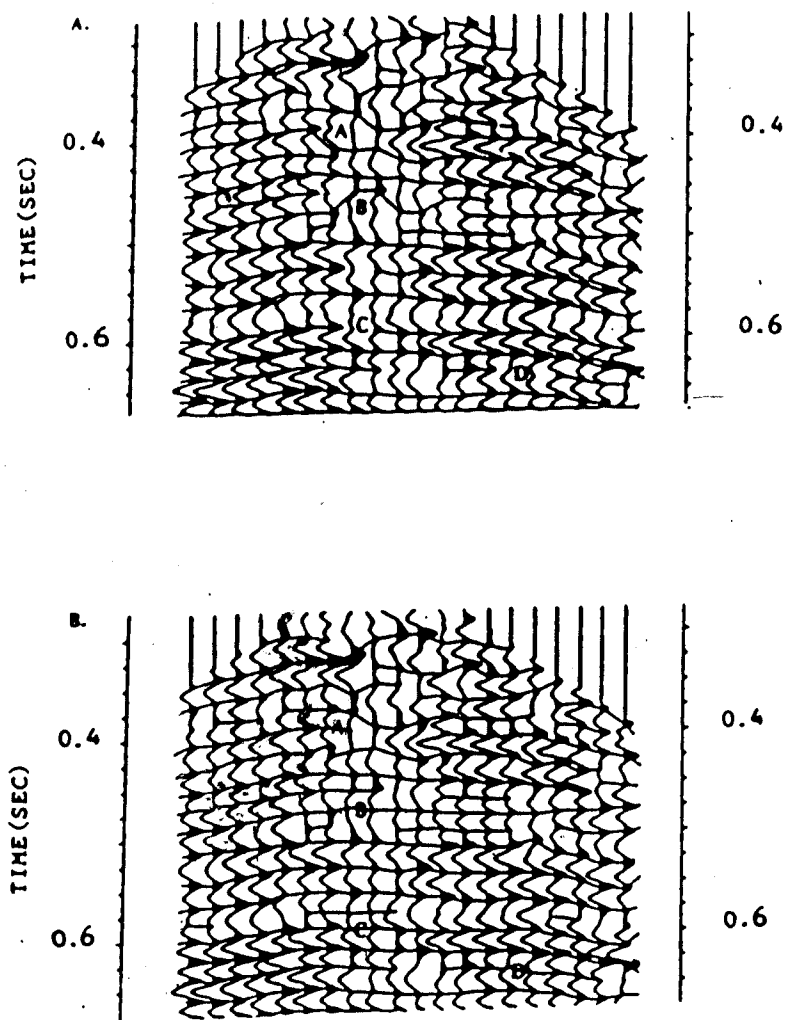


Figure 30.... A comparison of Lu's scheme (A) with mine (B).

the local peaks in the negative portions of the trace appear, some of my correlations continue through and Lu's fail. The grand mean of the four terms of the cost functions are respectively, 0.55 for location, 0.34 for duration, 0.95 for amplitude and 0.11 for overlap. The sensitivity of the matching results has been studied by varying the weights. A variation of ± 0.2 of each weight does not affect the results significantly.

It is not difficult to pinpoint the mismatches in Figure 29. Figure 31, which is the output of the array R, shows the location of mismatches (say A, B, C, D and G) more clearly. Each line segment represents a detected event. Segments of highest confidence are shown in thick solid lines, while those of least confidence are in dotted lines. The mismatches do not follow the general trend of the surrounding segments. The problem is not one of statics, whose correction has not been applied to these data. They might be the correct matches in case of faulting. However, faulting is not expected in this particular record. Figure 32 is the output of the matrix, C. It indicates that the mismatches have small correlation confidence.

Figure 33 and Figure 34 are the correlation results after parallel analysis. The window applied is 5 traces wide and 0.2 seconds long. The matches at A, B and C (see Figure 34) have been improved correctly while the mismatch at G remains unchanged. Correction at D has been erroneously made since the 'improved' result does not conform to the

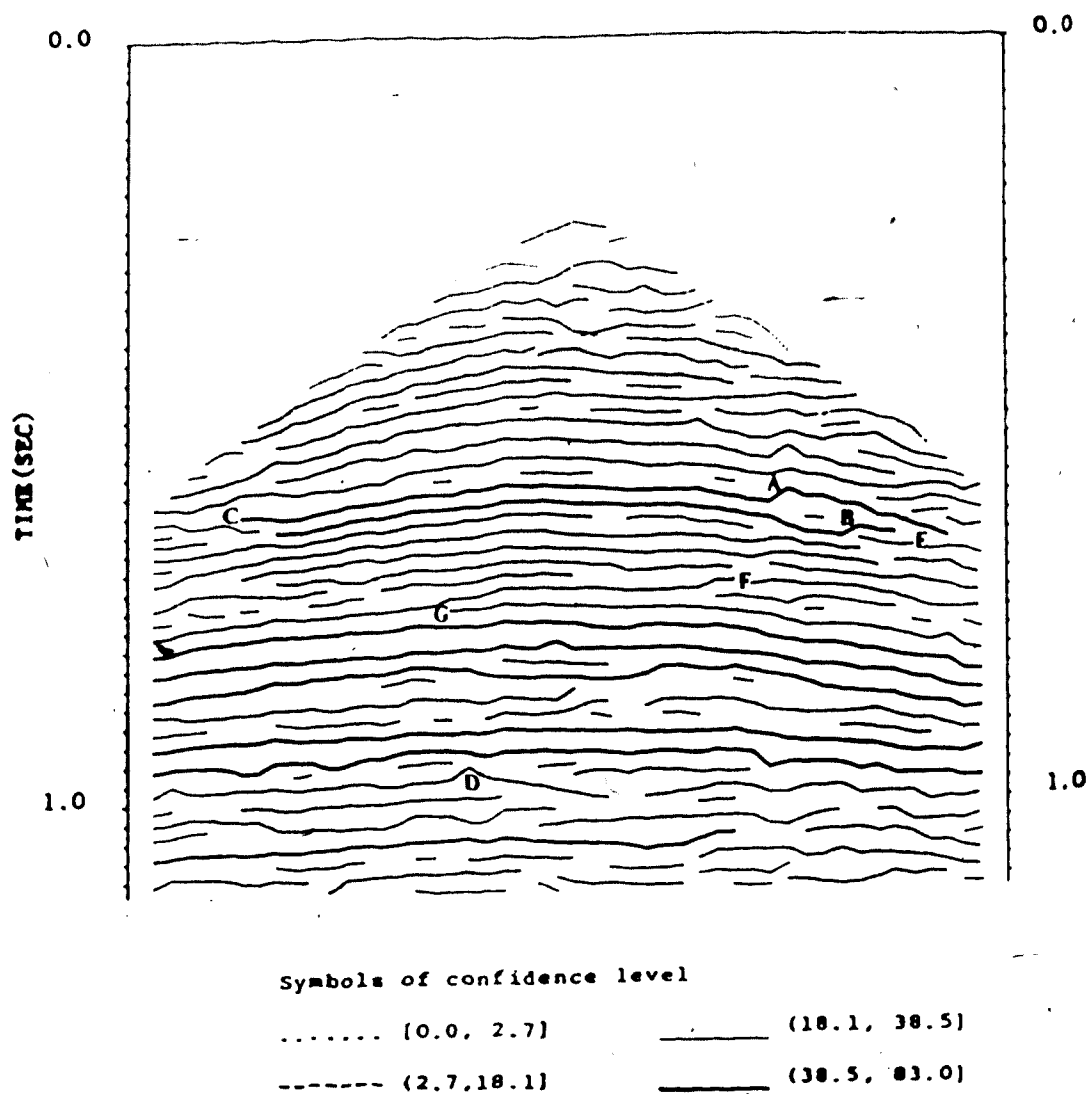


Figure 31.... Display shows the coherent segments in record 45 with coherence confidence.

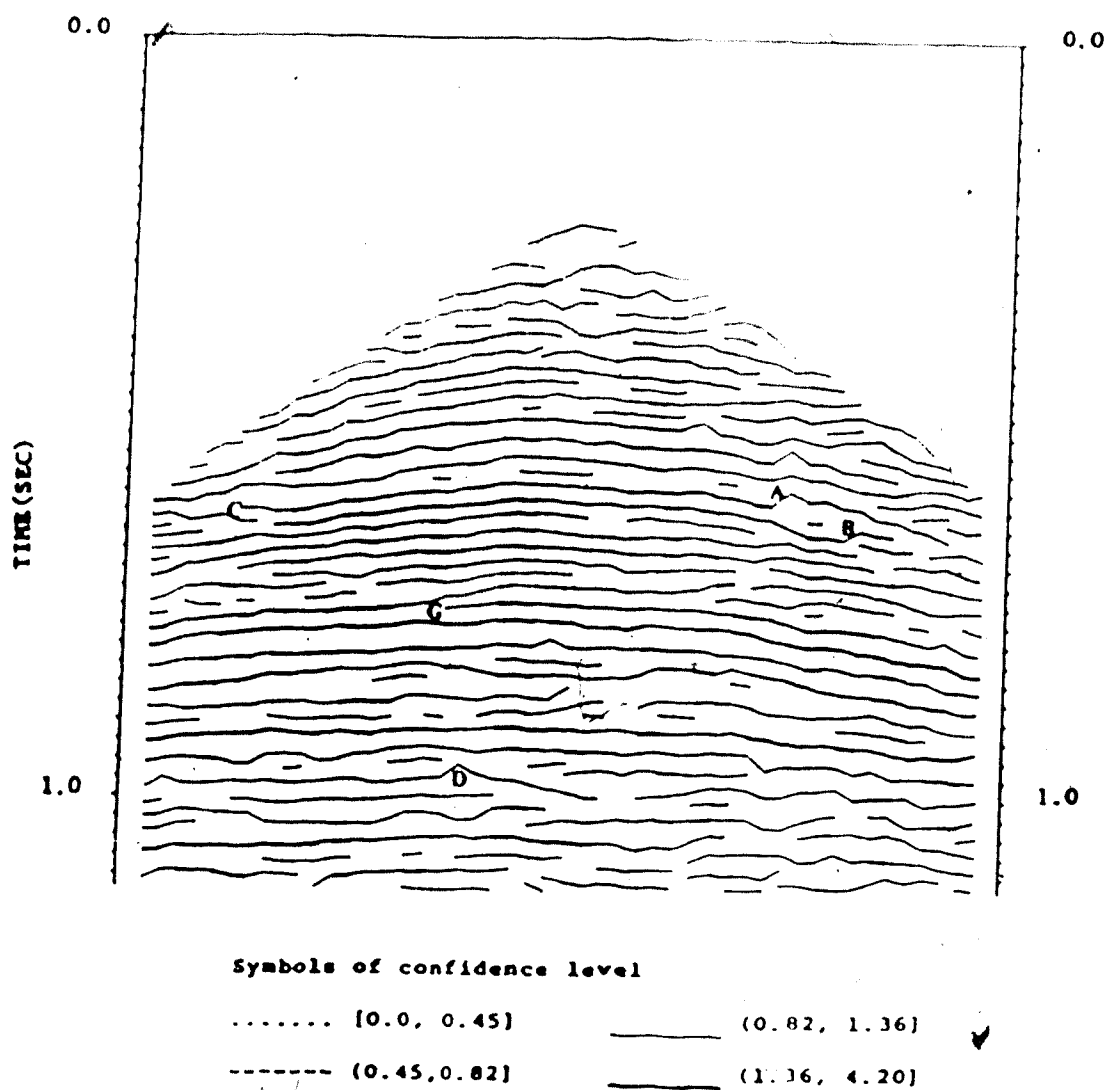


Figure 32.... Display shows the correlation segments in record 45 with correlation confidence.

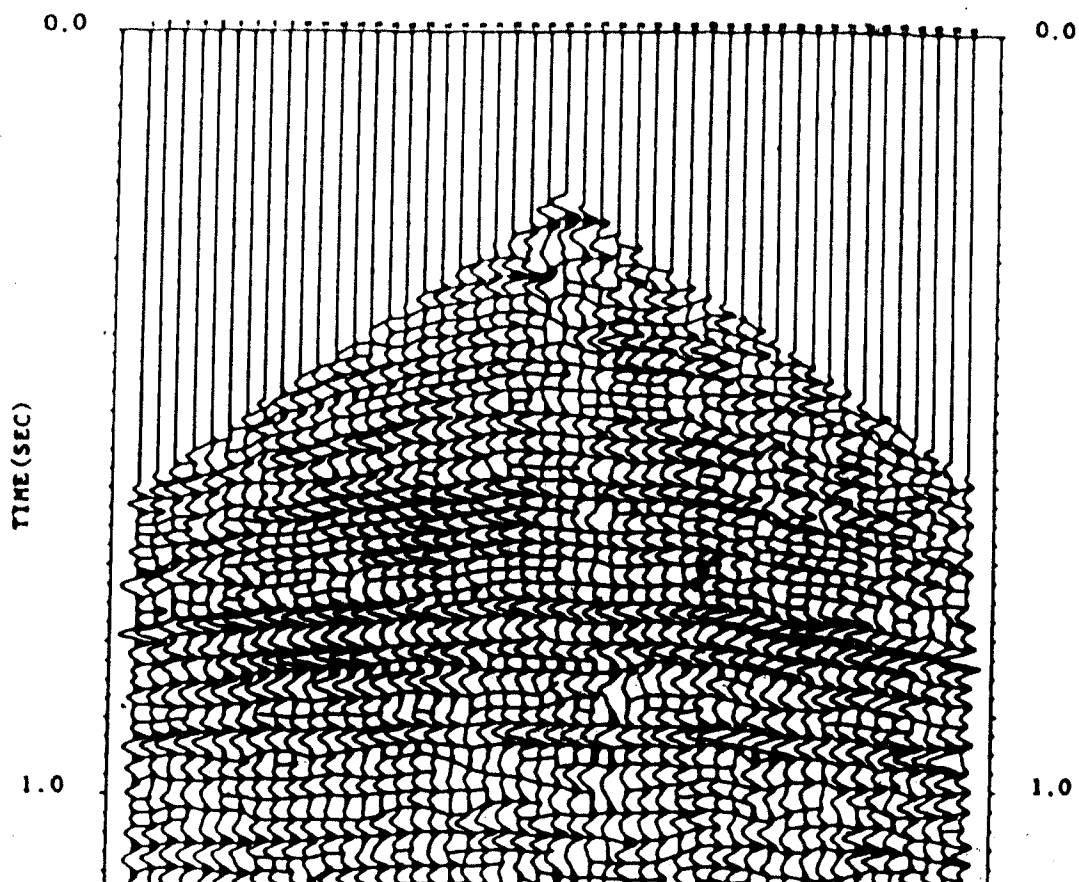


Figure 33.... Correlation result of record 45 after parallel analysis has been applied.

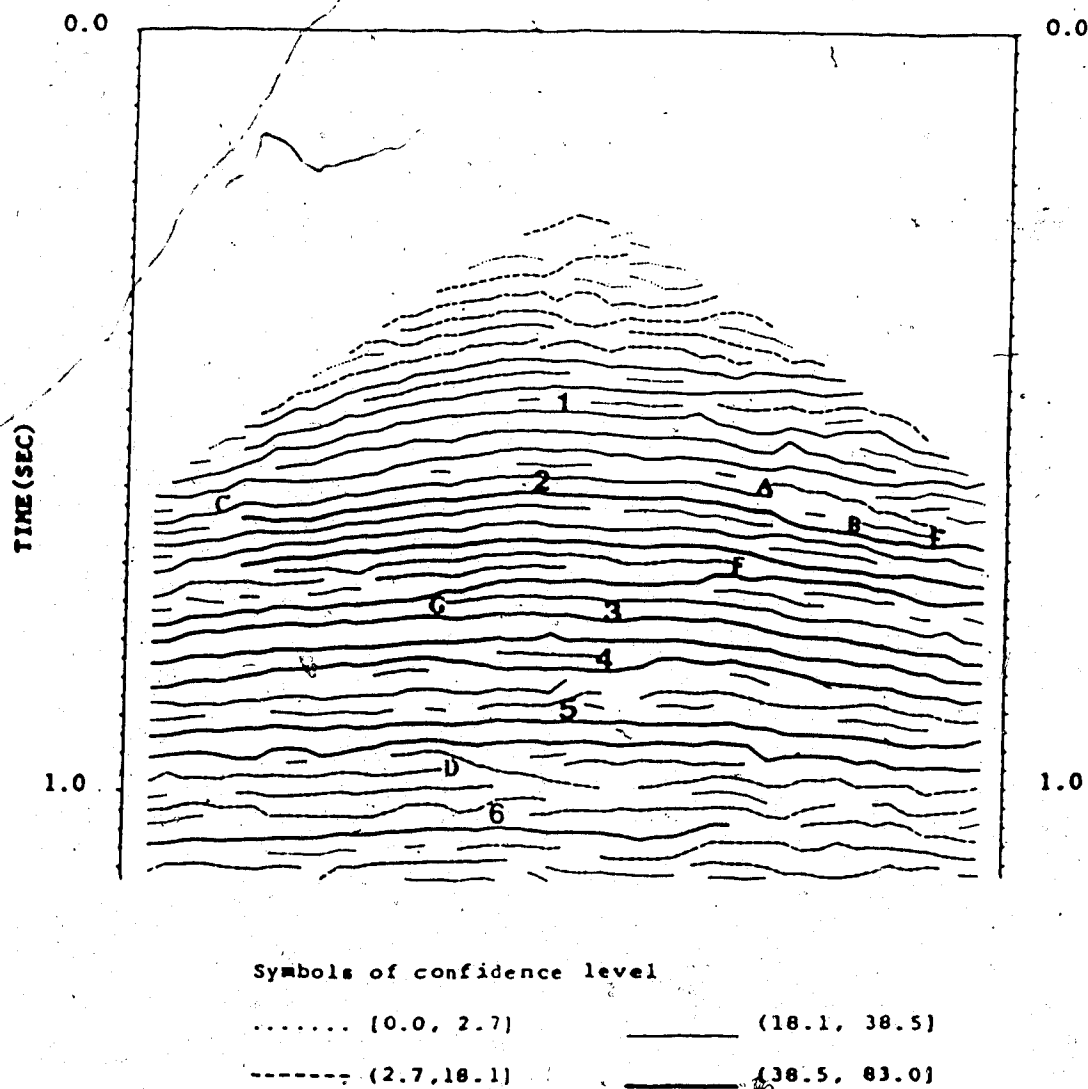


Figure 34.... Display shows the coherent segments in record 45 after parallel analysis has been applied.

surrounding events. Since the event is short and of low confidence, the error does not destroy the general performance of the scheme and thus can be ignored. Also, new pairs have been added at E and F. Joining at E has nicely completed the tracing of the major event. Joining at F is detrimental since the event indicated by the thick solid line is not a consistent coherent event as it appears to be. The joining has mistakenly raised the confidence of that event to the highest confidence bracket. It was expected that the mismatch at G could have been corrected, but criteria were not satisfied. The correction demands a further refining operation such as hyperbola-fitting. In the simple geology of the area under study, the events in a common shot gather are expected to follow hyperbolic trajectories. The trajectory fitting operation enlarges the correlation threshold to a segment. Breaking of segments will happen at points where the segments deviate from their corresponding hyperbola. It is also recommended that human interaction such as editing is profitable at this stage to do the correction. Figure 35, Figure 36, Figure 37 and Figure 38 show the correlation results of records 46 and 47 after parallel analysis.

Most of the major events identified in Figure 25, 26 and 27 have been recognized with high confidence. In record 45 (see figure 34), events 2, 3, 4, 5 and 6 are all recognized as events of highest confidence. Event 1 is recognized as an event in the second highest confidence

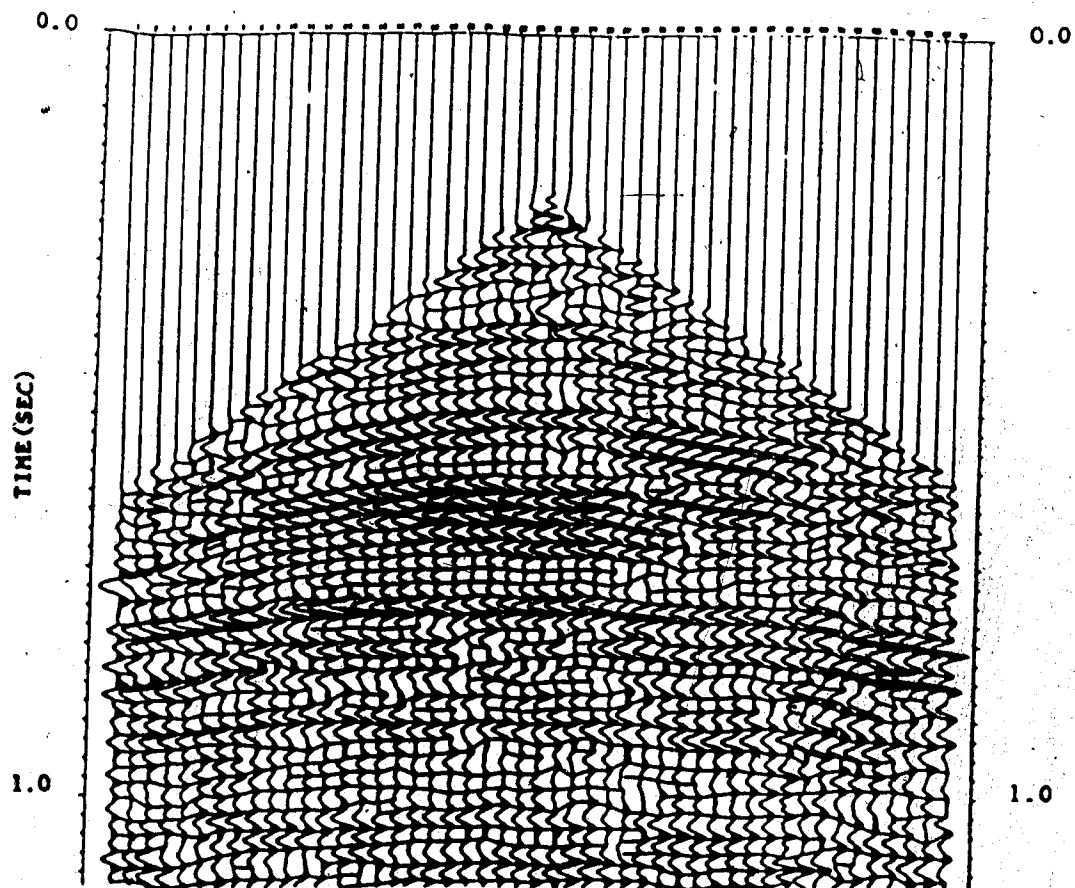
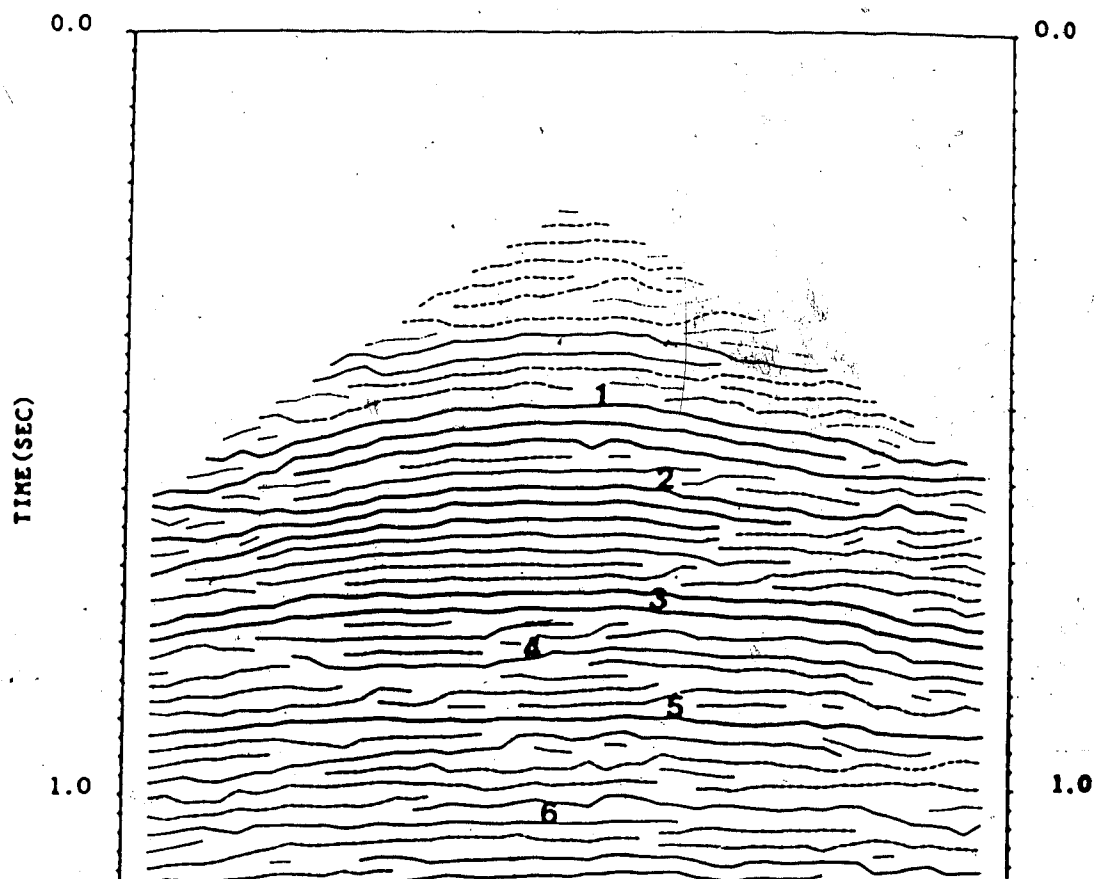


Figure 35.... Correlation result of record 46 after parallel analysis has been applied.



Symbols of confidence level

..... [0.0, 2.7]	_____ (18.1, 38.5)
----- (2.7, 18.1)	_____ (38.5, 83.0)

Figure 36.... Display shows the coherent segments in record 46 after parallel analysis has been applied.

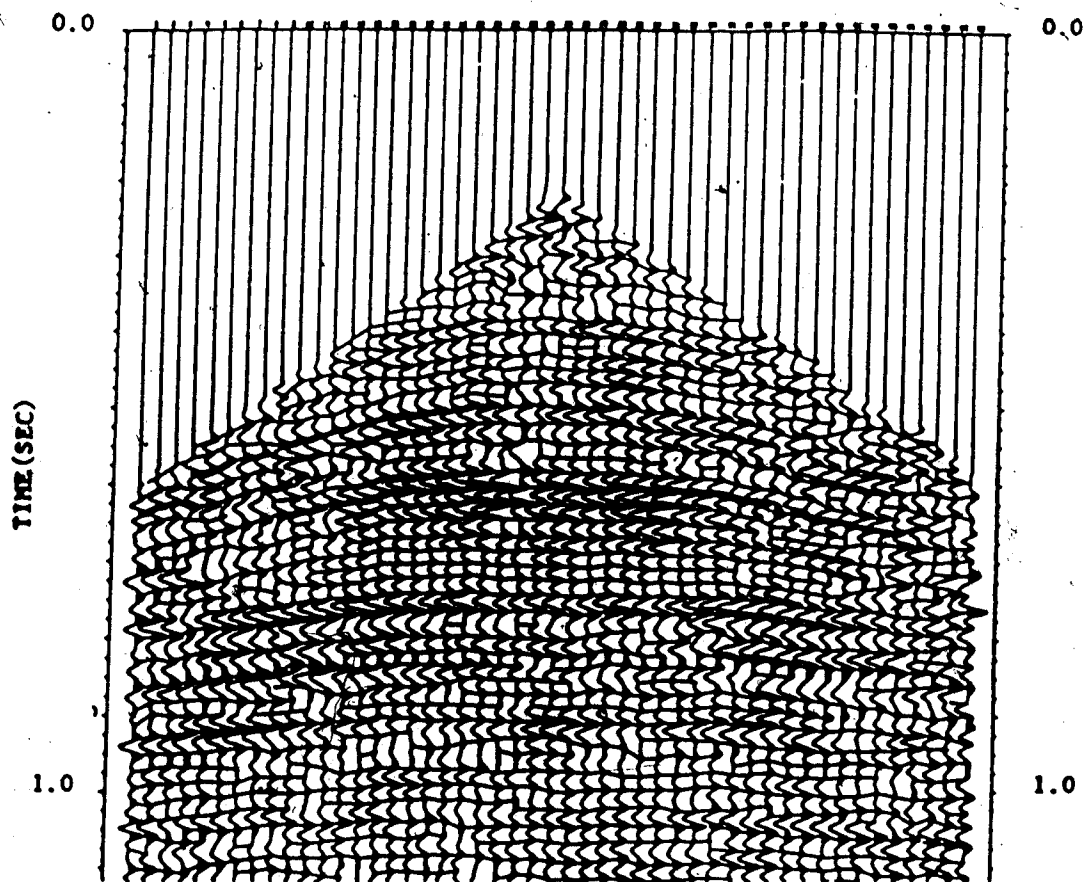


Figure 37.... Correlation result of record 47 after parallel analysis has been applied.

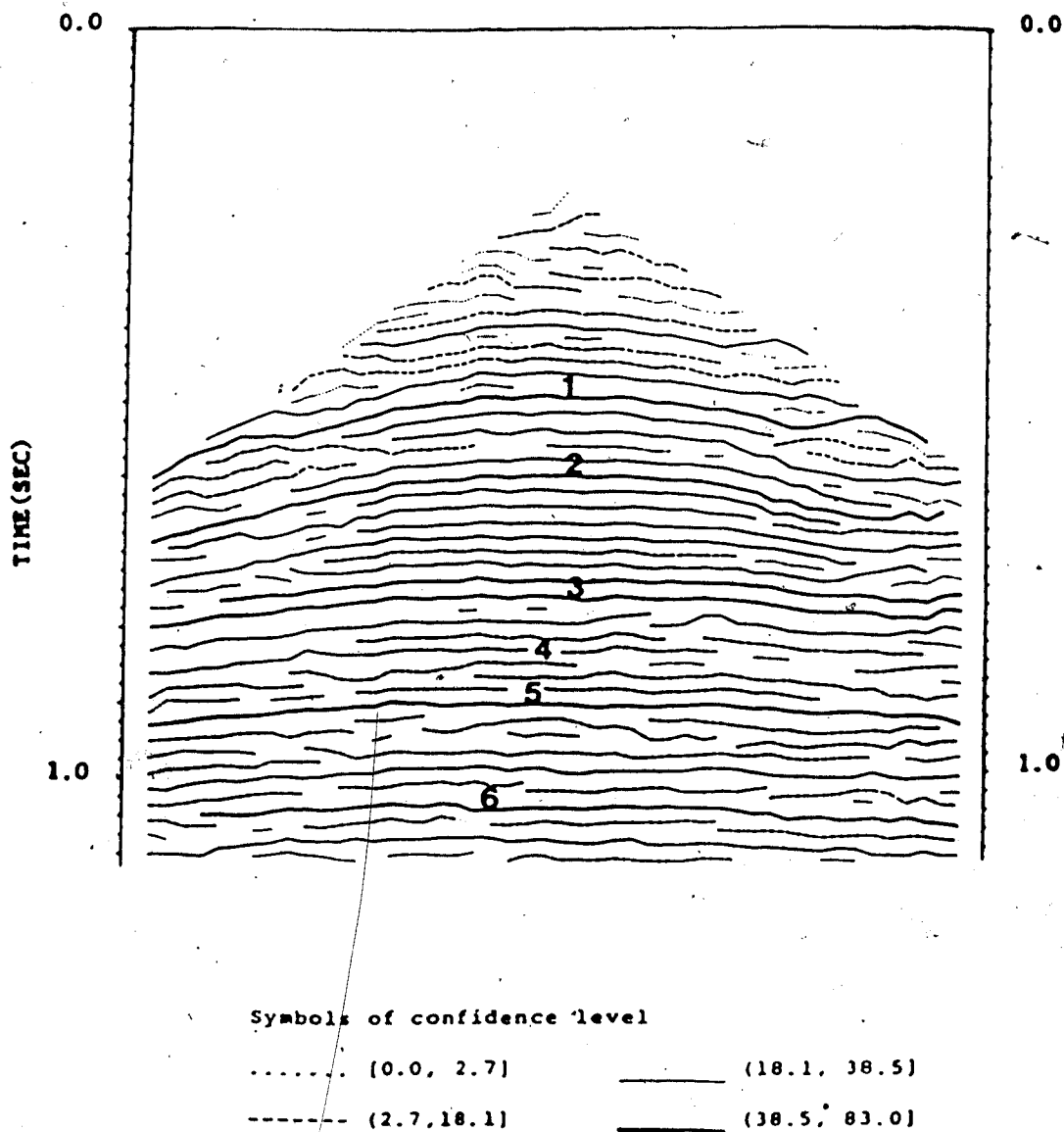


Figure 38..... Display shows the coherent segments in record 47 after parallel analysis has been applied.

level. This is due to the fact that there are some very small peak amplitudes associated with that event. Between trace 37 to 40, the costs of correlation are high. These two factors and another reason which will be discussed below contribute to the overall lower confidence of the whole segment. Nevertheless, extracting the event 1 is not a major problem due to the distribution of confidence values. Figure 39 shows the cumulative distribution of coherence confidence versus percentage of the segments detected in record 45. The segments have been sorted from low confidence to high confidence before plotting. It can be seen that the confidence values of the major events are quite discrete. This indicates that a decent adjustment of the confidence threshold level will pick up event 1 easily. In record 46 (see Figure 36), events 1, 2, 3 and 5 are recognized as events of highest confidence. Event 6 is nicely picked but classified to be in the second highest level of confidence. However, picking event 4 in this record has a bit of difficulty; indeed, some places in the segment have been incorrectly correlated. This is due to the fact that in some places, a waveform of single peak becomes one of double local peaks. In this case, the scheme traces the event inconsistently. In record 47 (see Figure 38), events 1, 2, 3, 5 and 6 are nicely picked. Because of some termination of the peaks along the event, event 4 has been recognized as a series of short segments; nevertheless, the event has been consistently traced.

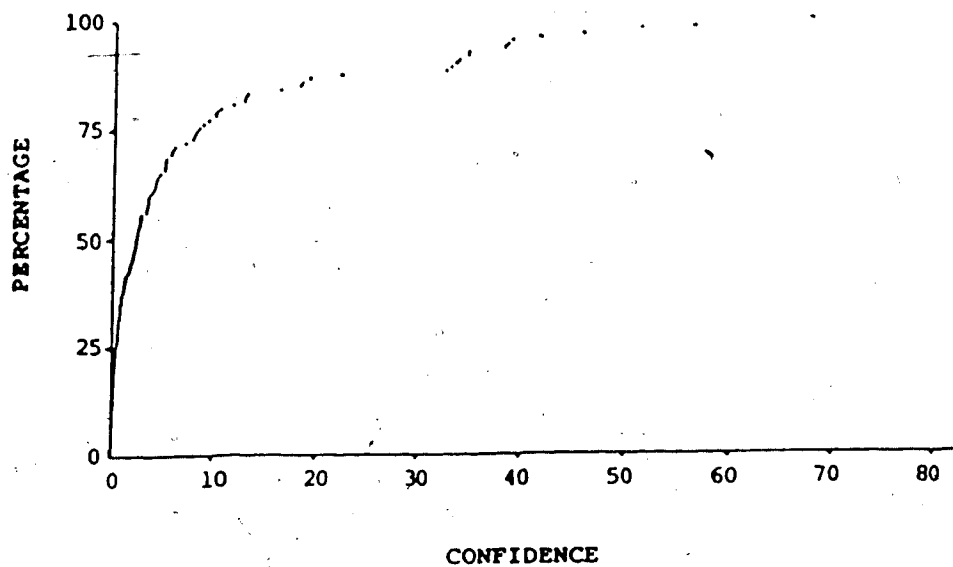


Figure 39.... Cumulative distribution of coherence confidence versus percentage of the segments detected in record 45.

Another problem can be seen obviously when the record 45 (see Figure 34) is analyzed. Some of the shallow reflections are good coherent events, but have lower coherence confidence values than those in the deep part of the record. This is due to the implicit assumption of the matching scheme. The location terms of the cost functions imply that the coherent events lie horizontally, thus attaching a high cost to those pairs which do not meet the requirement. This assumption works fine for a migrated section; however, in a common shot gather, the events display hyperbolic shape. Due to the greater moveout time, the shallow reflections have higher correlation costs than the deep ones. The higher correlation costs reduces the coherence confidence of the early events.

For record-to-record comparison, the record is divided into zones in terms of the coherent events which lie within the interval of highest confidence. Each record is then represented at location of trace 25 as a string of zones with features. Table 1, 2 and 3 are results of zonation in which record 45 is divided into 13 zones; record 46, 9 zones; record 47, 11 zones. Figure 40, Figure 41 and Figure 42 display the relative feature curves.

The zones are clearly divided into two groups with striking properties which can easily distinguish one from the other. Within each group, the properties are less distinguishable. There is a general trend of decreasing instantaneous frequency with arrival time. Based on this

NUMBER	LOCATION sec.	AV. LENGTH trace	AV. CONFIDENCE amp/cost	AVERAGE PEAK AMP	AV. COR. CONF. amp/cost	AV. INST. FREQUENCY	AVERAGE ENVELOPE
1	0.216, 0.61	8.5	2.2	1.2	0.4	38.3	1.2
2	0.61, 0.61	43.0	48.2	1.8	1.0	39.0	1.6
3	0.61, 0.67	20.0	5.5	1.1	0.7	39.8	1.1
4	0.67, 0.67	43.0	43.3	1.4	0.8	41.8	1.3
5	0.67, 0.73	11.5	5.4	0.5	0.3	36.9	1.1
6	0.73, 0.73	48.0	40.0	1.0	0.6	40.0	1.1
7	0.73, 0.77	3.7	0.7	0.5	0.2	37.7	1.3
8	0.77, 0.84	48.0	62.3	2.0	1.1	36.4	1.4
9	0.84, 0.91	4.9	0.8	0.7	0.3	31.0	1.2
10	0.91, 0.94	48.0	51.1	1.7	1.1	32.4	1.5
11	0.94, 1.06	14.3	5.3	0.8	0.4	33.4	1.1
12	1.06, 1.06	34.0	41.2	1.7	1.2	29.9	1.4
13	1.06, N/A	9.1	4.0	0.9	0.4	32.0	1.1

Table 1... Zones' features of shot 25.

NUMBER	LOCATION sec.	AV. LENGTH trace	AV. CONFIDENCE amp/cost	AVERAGE PEAK AMP	AV. COR. CONF. amp/cost	AV. INST. FREQUENCY	AVERAGE ENVELOPE
1	0.216, 0.50	8.4	2.7	1.1	0.4	43.2	1.1
2	0.50, 0.54	32.0	18.0	1.7	0.9	40.9	1.4
3	0.54, 0.60	7.4	2.9	0.8	0.4	48.2	1.2
4	0.60, 0.64	15.8	51.6	1.8	0.9	44.3	1.2
5	0.64, 0.74	19.5	8.3	0.9	0.6	44.9	1.1
6	0.74, 0.76	48.0	69.5	2.1	1.4	40.3	1.4
7	0.76, 0.90	10.4	3.7	1.1	0.6	35.8	1.2
8	0.90, 0.90	48.0	66.3	1.8	1.4	39.3	1.4
9	0.90, N/A	18.9	6.8	1.0	0.5	35.8	1.1

Table 2.... Zones' features of shot 46.

NUMBER	LOCATION sec.	AV. LENGTH trace	AV. CONFIDENCE amp/cost	AVERAGE PEAK AMP	AV. COR. CONF. amp/cost	AV. INST. FREQUENCY	AVERAGE ENVELOPE
1	0.216, 0.50	6.2	1.6	1.0	0.4	42.2	1.1
2	0.50, 0.50	46.0	39.9	1.6	0.8	39.3	1.4
3	0.50, 0.60	7.6	3.6	1.4	0.6	41.4	1.3
4	0.60, 0.60	47.0	72.2	2.4	1.2	41.0	1.7
5	0.60, 0.74	16.9	5.1	1.1	0.6	44.5	1.2
6	0.74, 0.77	46.0	62.5	2.0	1.3	39.7	1.5
7	0.77, 0.91	10.7	4.2	1.1	0.6	35.0	1.2
8	0.91, 0.91	48.0	57.8	1.6	1.0	40.0	1.4
9	0.91, 1.05	10.4	3.7	0.8	0.4	35.7	1.1
10	1.05, 1.05	45.0	47.5	0.6	0.9	33.8	1.4
11	1.05, N/A	10.1	4.0	0.8	0.5	35.6	1.1

Table 3... Zones' features of shot 47.

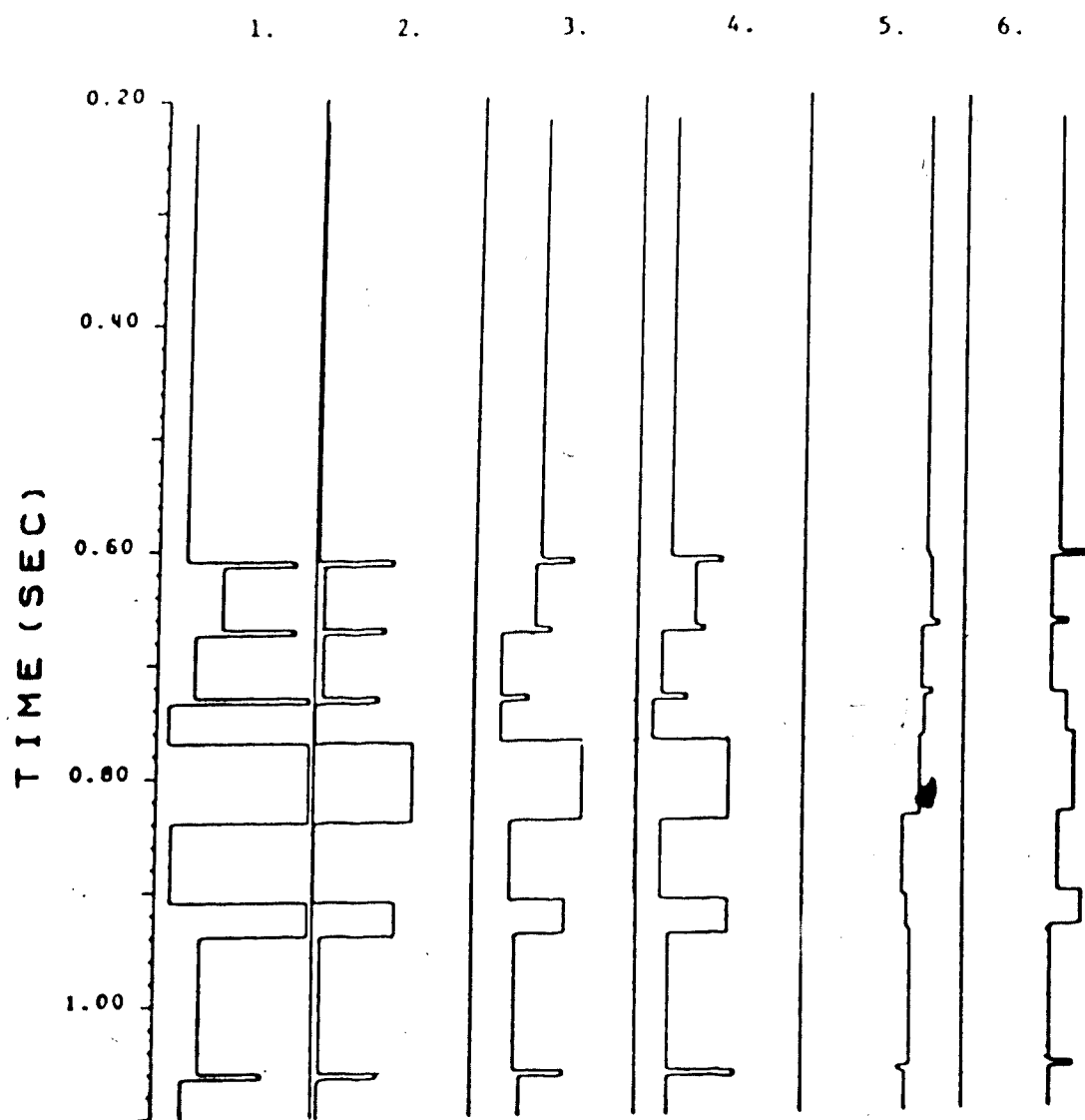


Figure 40.... Curves of zone features of record 45:
1=length; 2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope.

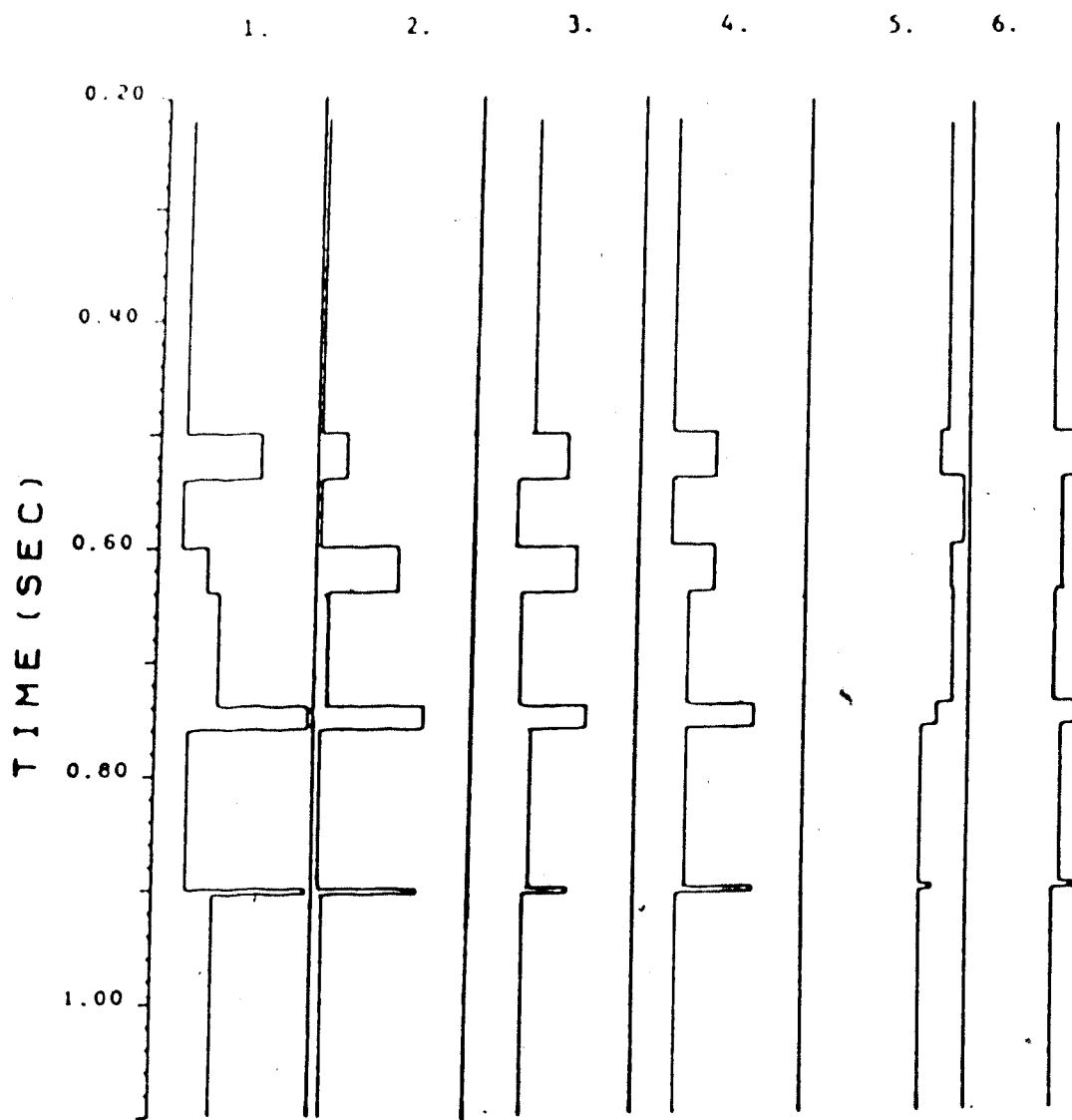


Figure 41.... Curves of zone features of record 46:
1=length; 2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope.

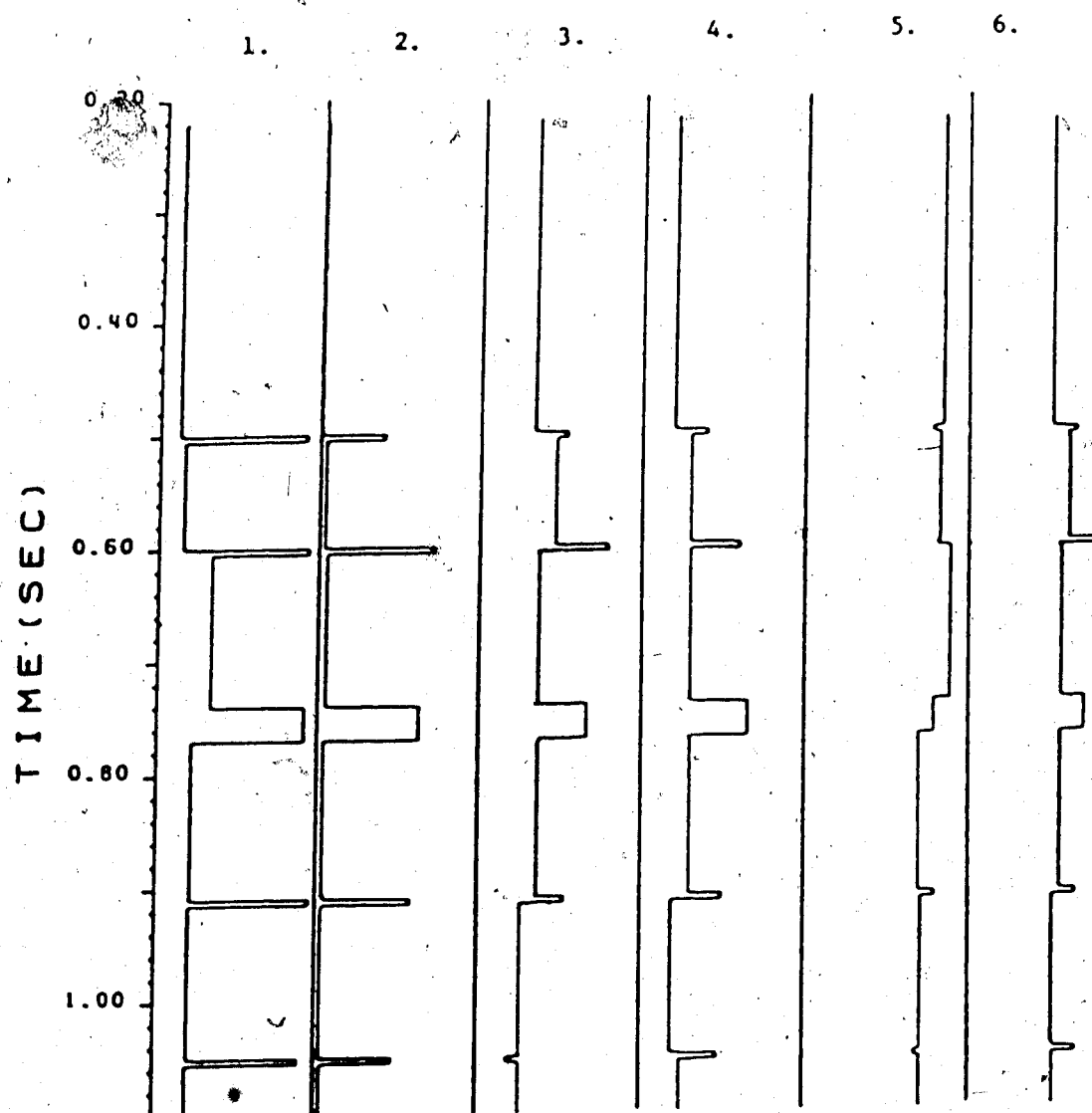


Figure 42.... Curves of zone features of record 47:
1=length; 2=coherence confidence; 3=amplitude; 4=correlation
confidence; 5=instantaneous frequency; 6=envelope.

background trend, zone 9 of record 45 has a lower frequency than the deeper part. This character can be traced to zone 7 in record 46 and zone 7 in record 47. This property is helpful to fix the reference in zone identification.

The zones in three records show good correlation in zone. Zones 2, 4, 6 and 8 in record 46 can be easily identified as zones 2, 4, 6 and 8 in record 47. Correlation has some difficulty in record 45 because some events have been erroneously recognized as major events; nevertheless, zones 2, 8, 10 and 12 correlate with zones 4, 6, 8 and 10 in record 47. The correlation is best suited to the string-to-string matching approach in which zones can be absorbed into others or deleted if they are misleading.

The quality of zonation greatly depends upon the result of the major event detection. Up to this point, it can be said that some zones are the approximate representation of the formation, bounded by two major reflection events. Without more information, no further conclusion can be made.

5. DISCUSSION AND CONCLUSION

There is no doubt that the local model-driven algorithm will be the dominant algorithm of seismic processing. Local geology and external information other than seismic data are crucial to the construction of the models which guide the processing and interpretation; however, seismic records themselves should be a major contribution. Also, iterative processing would be the most common routine, since modification of the models based on previous results is mandatory to better and more precise models corresponding to the real Earth's structure. Building up those models requires the picking of the major reflection events from the seismic records. Therefore, event recognition and record analysis based on the recognized events are two of the most fundamental operations in seismic interpretation.

The automated seismic record analysis described in this thesis attempts to emulate the procedure of visual analysis. The application has demonstrated AI's potential to bring the picking routine to automation once thought beyond computer capacity. Although the algorithm described is not typically knowledge-based, it can be used as a framework upon which more knowledge and rules can be accumulated. The success of such a computerized analysis can not only raise the level and productivity of data processing and interpretation, but also aid explorationists in decision-making. Such an automated system can be made even more efficient with the editing mode or interaction of human interpreters. The

conclusions in this thesis can be summarized as follows:

- a. Pattern recognition techniques are a promising tool in seismic record analysis and interpretation. The task can be profitably approached by three subtasks: trace-to-trace correlation, coherence analysis and zone analysis.
- b. In such an automated correlation procedure, dynamic matching such as a string-to-string matching algorithm has more advantages than the conventional cross correlation. Primitive selection is an important step. Ignoring the local negative peak has no geophysical justification; in fact, including them has yielded much better results. The cost functions used satisfy the classical distance definition and embody human experience in recognition.
- c. Confidence measure is an indication of the uncertainty of automated decision. Their introduction is necessary to understand the judgement made by the algorithm. With the confidence attached to the recognized events, the major events can be easily extracted for further analysis.
- d. Refining and iterative operations such as parallel analysis can improve the algorithm's performance. These operations usually consider more external information for decision. It is also recommended that the fitting of hyperbolae to the segments in a

shot gather can break the segments at points where they deviate far from the hyperbolic trajectories. The hyperbolae thus derived yield the image of the reflectors and the velocity profiles of the record.

- e. Record-to-record correlation can be made by reducing each record into a string of zones with features, thus yielding a new way to do the formation identification. Zone correlation can be done using the string-to-string matching algorithm. Six zone features have been extracted from each record. Two distinguishable groups have been obtained: one is of high confidence and the other is of low confidence. The two groups have striking difference in every aspect. From the analysis of the three records, complex attributes, coherence confidence, correlation confidence and amplitude have been identified as good discriminating features.

BIBLIOGRAPHY

- Aboutajdine, D., and Najim, M., 1981, A new approach for detection in explosive seismology: IEEE 1981 Computer-aided seismic analysis & discrimination, Aug., 19-21, North Dartmouth Mass.
- Aminzadeh, F., and Chatterjee, S., 1984, Applications of clustering in Exploration Seismology: Geoexploration, vol. 23, pp. 147-159.
- Anderson, K.R., and Gaby, J.E., 1983, Dynamic waveform matching: Computer-aided seismic analysis and discrimination, IEEE International Conference, pp. 98-108.
- Bois, P., 1980, Autoregressive pattern recognition applied to the delimitation of oil and gas reservoirs: Geophysical prospecting, vol. 28, pp. 572-591.
- Bois, P., 1981a, Reservoir recognition in petroleum prespection considered as an application of close man-machine communication: Proc, Symp. Computer Aided Analysis Discrimination, pp. 42-47.
- Bois, P., 1981b, Determination of the nature of reservoirs by use of pattern recognition with prior learning: Geophysical prospecting, vol. 29, pp. 687-701.
- Bois, P., 1982, Some comments on the application of pattern recognition to oil and gas exploration: Geoexploration, vol. 20, pp. 147-159.
- Bracewell, R.N., 1965, The Fourier Transform and its application: McGraw-Hill, N.Y., pp. 268-271.
- Diday, E. and Simon, J.C., 1976, Clustering analysis: in Digital pattern recognition, K.S. Fu Ed., Springer-Verlag Press, pp. 47-94.
- Farnbach, J.S., 1975, The complex envelope in seismic signal analysis: Bulletin of the Seismological Society of America, vol. 65, No. 4, pp. 951-962.
- Fu, K.S., 1974, Syntactic methods in pattern recognition: Academic Press.
- Fu, K.S., and Lu, S.Y., 1977, A clustering procedure for syntactic patterns: IEEE Trans. on SMC, SMC-7, no. 10, October.
- Gaby, J.E., and Anderson, K.R., 1984, Hierarchical

- segmentation of seismic waveforms using affinity: *Geop Exploration*, 23, pp.1-16.
- Gendzwill, D.J., 1978, Winnipegosis mounds and Prairie formation of Saskatchewan--seismic study: *The American Association of Petroleum Geologists Bulletin*, vol. 62, pp.73-86.
- Hagen, D.C., 1981, The application of principal components analysis to seismic data sets : 1981 2nd International Symposium on Computer Aided Seismic Analysis and Discrimination, pp.98-109.
- Hagen, D.C., 1982, The application of principal components analysis to seismic data sets: *Geop Exploration*, 20, pp.93-111.
- Hoyle, I.B., 1986, Computer techniques for well-correlation: *Geophysical Prospecting*, 34, pp.648-664.
- Huang, K.Y., and Fu, K.S., 1983, Classification of Ricker wavelets and the detection of bright spots using a tree classifier: 1983 3rd International Symposium on Computer Aided Seismic Analysis and Discrimination, pp.89-97.
- Huang, K.Y., and Fu, K.S., 1984, Detection of bright spots in seismic signals using tree classifiers: *Geop Exploration*, 23, pp.121-145.
- Kanasewich, E.R., 1981, Time sequence analysis in Geophysics: 3rd. edition, University of Alberta Press, chapter 21.
- Kirilin, R.L., Hutchins, R., Cudzilo, B., Dewey, L.A., and Hailey, L.L., 1984, The enhancement of seismogram parameters using image processing techniques: *Geop Exploration*, 23, pp.41-76.
- Lindseth, R.O., 1982, Digital processing of geophysical data: a review: Continuing Education Program, SEG.
- Liu, Hsi-Ho, and Fu, K.S., 1982a, A syntactic pattern recognition approach to seismic discrimination: *Geop Exploration*, 20, pp.183-196.
- Liu, Hsi-Ho, and Fu, K.S., 1982b, A syntactic approach to seismic pattern recognition: *IEEE Trans. on PAMI*, PAMI-4, no.2, March, pp.136-140.
- Lu, S.Y., and Fu, K.S., 1978, A sentence-to-sentence clustering procedure for pattern analysis: *IEEE Trans. on SMC*, SMC-8, no. 5, May.
- Lu, S.Y., 1982, A string-to-string correlation algorithm for image skeletonization: *Proceedings of the 6th*

International Joint Conference on Pattern Recognition,
Oct. 19-22.

Matlock, R.J., and Asimakopoulos, G.T., 1986, Can seismic stratigraphy problems be solved using automated pattern recognition?: Geophysics: The Leading Edge of Exploration, September, pp.51-55.

Oldenburg, D.W., Levy, S., and Whittall, K.P., 1981, Wavelet estimation and deconvolution: Geophysics, vol.46, pp.1528-1542.

Palaz, I., 1986, Artificial intelligence: expert systems for exploration: Geophysics: The Leading Edge, pp.60-64.

Pavlidis, T., 1977, Structural pattern recognition: Springer-Verlag Press.

Paulson, K.V., and Merdler, S.C., 1968, Automatic Seismic Reflection Picking: Geophysics, vol. 33, pp.431-440.

Robertson, J.D., and Nogami, H.H., 1984, Complex seismic trace analysis of thin beds: Geophysics, vol.49, no.4, pp.344-352.

Robinson, E.A., 1985, Digital seismic inverse methods: IHRDC, Boston, chapter 4.

Robinson, J.E., 1978, Pitfalls in automatic lithostratigraphic correlation: Computers & Geosciences, vol.4, no.3, pp.273-275.

Schneider, W.A., 1971, Developments in seismic data processing and analysis: Geophysics, vol.36, pp.1043-1073.

Schultz, P.S., 1985, Seismic data processing: Current industry practice and new direction: Geophysics, vol. 50, pp.2452-2457.

Shaw, B.R., and Cubitt, J.M., 1979, Stratigraphic correlation of well logs: an automated approach: in "Geomathematical and petrophysical studies in sedimentology", Computer & Geology, vol.3, edited by Gill, O., and Merriam, O.F.

Sinvhal, A., and Khattri, K., 1983, Application of seismic reflection data to discriminate subsurface lithostratigraphy: Geophysics, vol.48, pp.1498-1513.

Strom, T., 1976, On amplitude-weighted instantaneous frequencies: IEEE Trans. on ASSP., vol.25, pp.351-353.

Taner, M.T., and Sheriff, R.E., 1977, Application of amplitude, frequency, and other attributes to

stratigraphic and hydrocarbon determination: in Applications to hydrocarbon exploration, C.E. Payton, Ed., AAPG Memoir 26: Tulsa, Am. Assn. Petroleum Geologists, pp.301-327.

Taner, M.T., Koehler, F., and Sheriff, R.E., 1979, Complex seismic trace analysis: Geophysics, 44, pp.1041-1063.

Tjostheim, D., 1977, Recognition of waveforms using autoregressive feature extraction: IEEE Trans. on C, C-26, pp.268-270.

Wagner, R.A., and Fisher, M.M., 1974, The string-to-string correction problem: J. ACM, vol.21, Jan.

Waterman, D.A., 1986, A guide to expert systems: Addison Wesley Press.

Wu, X., and Nyland, E., 1985, The design and implementation of a computerized well log interpretation system:

Wu, X., and Nyland, E., 1986, Well log interpretation using Artificial Intelligence technique: SPWLA Twenty-Seventh Annual Logging Symposium, June 9-13, Houston, U.S.A.

Ziemer, R.E., and Tranter, W.H., 1976, Principles of communications: systems, modulation, and noise: Houghton Mifflin Company Press, pp.72-79.