PATTERN RECOGNITION FOR EARTHQUAKE DETECTION

By Manfred Joswig

Abstract

The detector algorithms in use at date rely on negative decision logic: based on a model of the ambient noise process they detect all deviations, but many of them are false alarms. The principal alternative to this approach is pattern recognition, which tests on positive correlation with some known signal patterns.

The Sonogram-detector realizes this scheme for single seismogram traces. Sonograms display spectral energy versus time. Suitably scaled, these images display only information which is significant to the detection process. Patterns of known earthquakes and noise signals are defined by means of these images. Event detection is performed by recognizing one of the patterns in the actual sonogram. The overall processing scheme is similar to the visual inspection of seismograms by the human observer.

An off-line test installation for detecting local earthquakes proves the expected low false alarm rate, high timing accuracy and good detection probability of the Sonogram-detector.

Introduction

For more than 25 years automatic detection of earthquakes in environmental noise has been the goal of research efforts. Originally motivated by the large amount of on-line data provided by seismic arrays and regional networks, standard detectors like STA/LTA (Short Term Average to Long Term Average) are now well known in all fields of geophysical data acquisition.

In seismology none of the standard detectors reliably rejects false alarms produced by noise fluctuations. Detections shorter than a given time window, however, can be associated with noise pulses and then be ignored. Allen (1978, 1982) for the first time treated this logic as a separate *post-detection* process and introduced new criteria that are more complex. He also suggested the term *phase picker* for those specialized detectors, whose inherent timing inaccuracy is very low at the cost of the signal averaging length.

The STA/LTA-based detectors approximate the Neyman-Pearson filter and use signal averaging in the STA to archive an optimum S/N ratio (see Fig. 1). Another principal approach was introduced by Stewart (1977). He tunes signal processing for optimum postdetection criteria by introducing an arbitrary *characteristic function* (CF). But calculating statistics and decision on the CF is like applying to the seismogram any other filter but the optimum Neyman-Pearson one and will perform worse (Blandford, 1982). So CF-based detectors have their better detection to false alarm ratio only due to the significant power in their postdetection logic.

Both STA/LTA- and CF-based detectors model the ambient seismic noise level by some kind of mean and deviation. This noise model must be determined for every new installation and limits the detection threshold (see Fig. 1). Each pulse with energy above the threshold will trigger a (preliminary) detection, so these detectors inherently follow *negative* decision logic: The defined state (by mean and deviation) is noise or *nondetection.* Every change must be detected, since it might be induced by an earthquake. The preliminary detection may be evaluated by some postdetection logic. This test on positive agreement with a few simple rules can reduce the false alarm rate significantly.

	processing of seismogram			statistics of seismic noise		decision	postdetection
	filter x → y	characteristic function CF	energy during length of signal	mean	deviation		
Freiberger (1963)	$\sqrt{\dfrac{S}{N(S+N)}}$ S.N: PSD's		$STA = \dfrac{1}{n}\sum_{i=1}^{n} y_i^2$			$STA > \beta$	
STA/LTA Vanderk. et al 1965	$\dfrac{BP}{\sqrt{N}}$ BP bandpass		$STA = \dfrac{1}{n}\sum_{i=1}^{n} \lvert y_i\rvert$	$LTA_i = \alpha\,STA_i + (1-\alpha)\,LTA_{i-1}$		$STA > \beta\,LTA$	window
Z-Detector Swindell+ Snell 1977	BP		$STA = \dfrac{1}{n}\sum_{i=1}^{n} y_i^2$	$LTA_i = \alpha\,STA_i + (1-\alpha)\,LTA_{i-1}$	$\sigma_i^2 = \alpha\,(STA_i - LTA_i)^2 + (1-\alpha)\,\sigma_{i-1}^2$	$STA > LTA + \beta\,\sigma$	3 windows
WALSH Goforth+ Herrin'81	$\dfrac{BP}{\sqrt{N}}$		$M = \sum_{i\in q} \lvert Y_i\rvert$ window average	M_{50} percentiles from	$V = M_{75} - M_{50}$ reorder M by value	$M > M_{50} + \beta\,V$	window
Allen (1978)		$CF = x^2 + \gamma\,\dot{x}^2$	$STA_i = \delta\,CF_i + (1-\delta)\,STA_{i-1}$	$LTA_i = \alpha\,STA_i + (1-\alpha)\,LTA_{i-1}$		$STA > \beta\,LTA$	analysis of min-max-fct by adaptive window
SRO Murdock+ Hutt 1983	BP	$CF = \lvert\Delta pf(y)\rvert$ pdf: peak-trough			$s' = 2\,\sigma(CF)$ short term estimate	$CF > \beta\,s'$	window

Signal Processing in Traditional Single-Trace Detectors

FIG. 1. History of detector development starts by the theoretical approach of Freiberger, introducing both the Neyman-Pearson filter and the signal averaging for better S/N ratio. STA/LTA is the first implementation for practical use. The Z-detector is a refinement. Its detection threshold better adapts to the diurnal and seasonal changes in seismic noise level. The Walsh-detector performs the same way in sequency domain, so it is STA/LTA-based too. The CF-based detectors (Allen, SRO) optimize signal processing for the postdetection criteria. For more details and other detectors, see Joswig (1987).

PATTERN RECOGNITION DETECTORS

A completely different approach for detection is realized by knowledge-based reasoning. The *pattern recognition* (PR) implements a decision logic of positive kind: defined states are patterns of earthquakes and temporary noise signals, only a sufficient similarity with one of these patterns will trigger a detection.

Note here that the detection threshold or pattern fit is not dependent on the seismic noise level and that at this stage earthquakes and noise bursts are treated the same way. Detector installation, however, requires the definition of an appropriate pattern set.

According to the principles of *artificial intelligence* (AI), the signals are identified by inference rules acting on a set of knowledge patterns. So PR-detectors form a completely new class of detectors not related to the Freiberger-based characterization of traditional detectors in Figure 1.

One commonly agreed principle of AI is that choosing a specific knowledge representation determines the whole problem-solving approach. The most frequently used representations by frames, rules, or scripts (see, e.g., Winston, 1984)

are all propositional ones. When Kosslyn (1980) initiated the *imagery debate* in cognitive science, analog representations, (e.g., images) became the principal alternative. They are of first choice for all knowledge acquisition and application tasks, while propositional forms are suited best for knowledge reasoning (Haugeland, 1985).

Problem solving by mental images is simple, since mental patterns are as real to the mental images of sensing (i.e., the recognized seismogram) as visual patterns in a picture search. So the methods for pattern recognition are the same, chosen from the developed toolkit of computer vision techniques (see, e.g., Gonzales and Wintz, 1987).

The early PR-detectors in seismology apply some ideas of syntactic sentence analysis. Liu and Fu (1983) perform a cluster analysis on equidistant intervals in some given seismograms. The analysis is based on a few parameters like number of zero-crossings and maximum amplitude. The identification of each of these clusters with a letter transforms a seismogram into a sequence of letters. Defining a knowledge base of known sequences and an inference procedure like the *nearest neighborhood role* results in a complete detector, if a threshold for sufficient neighborhood is also implemented.

Other contributions to the idea of syntactic PR and its extension to a complete grammar of seismograms were supplied by Anderson (1981, 1982a, b), Gaby and Anderson (1984), Faure *et al.* (1984), and Chen (1984). All of these approaches tend to propositional knowledge representation, none of them uses the pictorial representation of seismograms. Of course, simply choosing selected seismogram plots as knowledge patterns will equal the design of matched filters; this approach must fail due to the unpredictable variety of individual seismograms. But human observers are able to detect and distinguish typical and simple forms of earthquakes even in very noisy environment. So their "processing scheme" needs some seismograms and intermediate, mental forms of pictorial knowledge representation.

Processing Steps of Pattern Recognition Detectors

The principal idea of PR on mental images developed so far yields to the following processing scheme: The detector transforms the measured seismograms into mental images; recognition as identification is performed by standard Computer Vision methods. So in detail the PR-detector consists at least of:

- The transformation of seismograms into mental images by nonlinear parametrization. This is like drawing a pen sketch. (Each task like detection, P- or S-onset picking will require another best suited mental image, derived from the same seismogram.)
- The knowledge base of typical earthquakes and noise bursts. This is a set of mental images. For a nonlearning algorithm they must be derived by the seismologist.
- The pattern recognition process. It acts on mental images and implements a detection threshold and a decision logic selecting the most similar pattern.

But why should noise bursts be characterized in the knowledge base, if one doesn't want them to be detected. The reason is inherent in testing on positive agreement done by all PR-detectors: An actual earthquake will never fit perfectly with one of the predefined patterns, so the detection threshold is always set below maximum.

A knowledge base without noise patterns requires a reasonably high threshold to avoid false alarms. With noise patterns, however, the threshold can be lowered significantly. Even if a noise burst triggers earthquake patterns now, the decision logic chooses the most similar one—most probably a noise pattern. This false alarm can be suppressed by the pattern identification, while the detection of earthquake patterns is unaffected. Overall the additional noise patterns give a better detection probability without an increased false alarm rate.

THE SONOGRAM-DETECTOR

The principal approach of PR on mental images was first realized for the detection of local earthquakes in the Bochum University Germany network (Joswig and Harjes, 1986; Joswig, 1987). The mental images are based on sonograms, i.e., pictures of spectral energy versus time. They are nonlinearly scaled, however, and differ significantly from spectral density plots published so far (Fleck and Fleck, 1964; Kennett, 1975; Sheppard, 1978; Evans and Allen, 1983).

Definition of the Transformation

The mental images of the Sonogram-detector are two-dimensional gray-scale pictures, where blackness is a measure of spectral energy.

To transform seismograms to mental images the power spectral density is calculated first. Data in the BUG-network are sampled at 100 Hz and divided in 2.56-sec intervals with 50 per cent coverage. Each interval is weighted by a cos-window and then Fourier-transformed. For parameter reduction the frequency axis is scaled logarithmically by averaging the power spectrum $a(f)$ in half-octave-wide passbands. Calculation for all time intervals gives the power spectral density matrix A by

$$A := \left\{ a_{ij} = a(f_i, t_j), \quad \begin{array}{l} f_i = 0.4, 0.6, 0.8, 1.1, 1.6, 2.1, 2.9, 4, 5.5, 8, 11, 16, 21\, Hz \\ t_j = t_0 + j\Delta t, \quad \Delta t = 1.25s \end{array} \right\}. \tag{1}$$

The mental images, however, are based on matrix B, which has been prewhitened by the vector $\{n_i\}$ containing the stationary noise power spectrum as

$$B := \left\{ b_{ij} = \frac{a_{ij}}{n_i} \right\}. \tag{2}$$

The second step for transforming seismograms into mental images is done by scaling B relative to a set of gray-scale thresholds g_k. The resulting picture matrix C is defined by

$$C := \{ c_{ij} = k \mid g_k \leq b_{ij} < g_{k+1}, \, k = 0, 1, 2, \ldots, 9 \}. \tag{3}$$

The contrast in C is sensitive to the distribution of scaling thresholds g_k over the dynamic range of PSD matrix B; some choices are shown in Figure 2, while Figure 3 gives the corresponding mental images. For *logarithmic absolute* scaling, the whole dynamic range 2^{2q}:1 is divided into equidistant steps by the thresholds

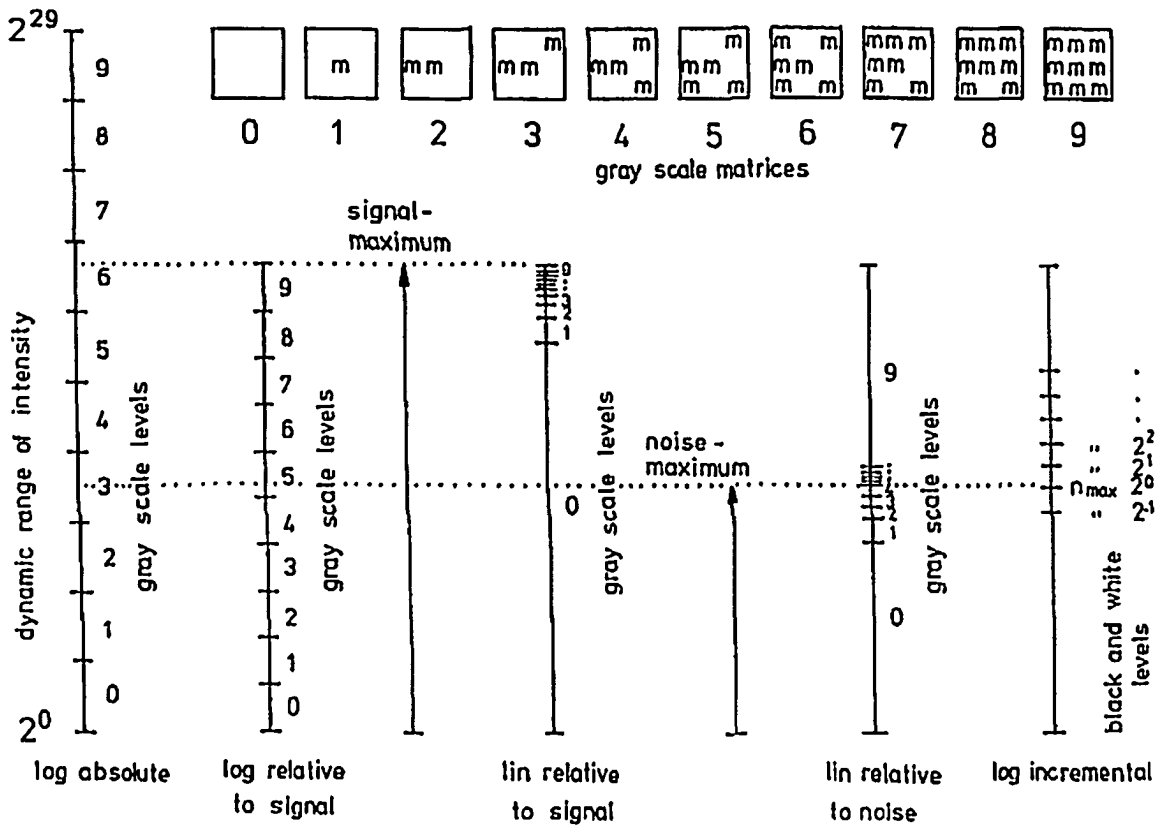$$g_0 = 0, \qquad g_k = 2^{2q(k/10)}. \tag{4}$$

FIG. 2. The mental images are composed by 3 * 3 elements of increasing blackness. Different scalings can be defined to map the dynamic range of matrix B in equation (2) to the 10 gray-scale levels. The effect of the first four scalings is shown in Figure 3. The log incremental scaling is used for the pattern recognition process in Figure 8.

The mental image has little contrast. This remains true for *scaling logarithmic relative* to signal maximum $b_{ij,max}$ by thresholds g_k'.

$$g_k' = g_k = \frac{b_{ij,max}}{2^{2q}}.$$ (5)

For *scaling linear relative* to signal maximum by

$$g_k = b_{ij,max} \frac{k}{10}$$ (6)

only the intensity maximum is visible. *Scaling linear relative* to the noise maximum $n_{i,max}$ gives the best result. It is used for the rest of this work. The mental image here corresponds quite well to the seismologist's impression of detectable signal energy in the seismogram. Scaling is done by

$$g_k = n_{i,max} \frac{k}{10} \frac{9}{n_{level}}, \qquad g_{10} = 2^{2q}.$$ (7)

The gray level of noise maximum $n_{i,max}$ has been set by convenience to $n_{level} = 4$, so all signal intensity above $g_9 \cong 2 * n_{i,max}$ will get black shading.
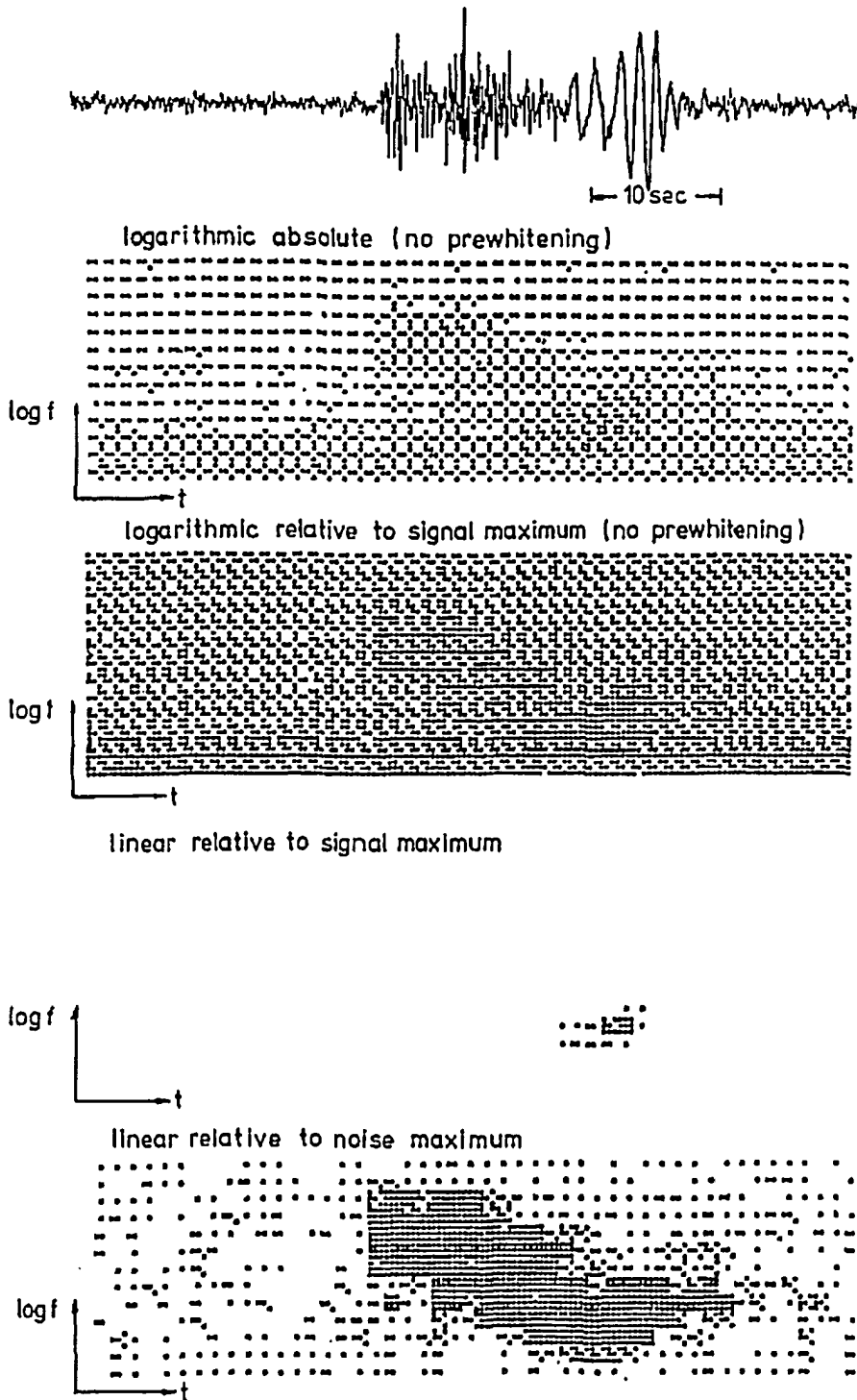
FIG. 3. One local event ($M_1 = 2.0$, $\Delta = 35$ km) is shown in mental images of different scaling (see Fig. 1). The two logarithmic scalings base on matrix A in equation (1) instead of B in equation (2). The best contrast results from the highly nonlinear scaling linear relative to noise.

One important task of the mental image defined so far is to perform an effective parameter extraction. A high resolution plot like the seismogram trace in Figure 3 has

$$2000 \text{ time samples} * 2^9 \text{ amplitude steps} = 1 * 10^6 \text{ bit}$$

information in the time versus amplitude plane. The equivalent mental image, containing the same information for the detection process, has less than 1 per cent, namely

48 time samples $*$ 13 frequency samples $*$ 10 gray steps $= 6 * 10^3$ bit.

Definition of the Knowledge Base

The knowledge base consists of a set of signal patterns. Each pattern represents an average sonogram, either from earthquake regions or from known noise burst types like sonic boom, traffic or industry noise from cars, trains, saw-mills etc. Figures 4 and 5 present sonograms of local earthquakes from two distinct epicentral distances; they average to one pattern each. Figure 6 gives several kinds of noise signals; they differ significantly from the earthquake patterns shown before.

An event pattern in the knowledge base is formally defined by matrix $\mathbf{M}$ and attached scaling factor N through

$$\mathbf{M} := \{m_{ij} \mid m_{ij} \in (-1, 0, +1)\} \tag{8}$$

$$N = \sum_i \sum_j \mid m_{ij} \mid . $$

To give an example the simple pattern set in Figure 7 consists of four local earthquake types and one teleseismic onset. For best match with an actual seismogram the striped areas with $m_{ij} = +1$ require signal energies above a suitable threshold, while the crossed areas marked by $m_{ij} = -1$ must have *no* energy. The remaining pattern with $m_{ij} = 0$ will be ignored in the PR-process. The edge conversion by $m_{ij} = -1$ greatly enhances the S/N ratio of pattern fit (12), if pattern $\mathbf{M}$ (9) approximately fulfils the balance condition.

$$\sum_i \sum_j m_{ij} = 0 \tag{9}$$

Definition of the Pattern Recognition Process

The PR-process for mental images is shown in Figure 8. Correlation between pattern $\mathbf{M}$ and the sonogram is much simplified when replacing its gray scale matrix $\mathbf{C}$ by a set of black and white pictures $\mathbf{D}_s$. As with $\mathbf{C}$ in (3), they are derived from PSD matrix $\mathbf{B}$ now by a set of increasing thresholds h_s as

$$\mathbf{D}_s := \left\{ d_{ij} = \begin{matrix} +1 \text{ for } b_{ij} > h_s \\ -1 \text{ for } b_{ij} \leq h_s \end{matrix} \right\}. \tag{10}$$

Thresholds h_s are scaled *logarithmic incremental* (see Fig. 2) from half of the noise maximum to signal maximum as

$$h_s = n_{i,\max} 2^s, \qquad s = -1, 0, 1, \ldots, \text{int}\left(\log_2 \frac{b_{ij,\max}}{n_{i,\max}} \right). \tag{11}$$
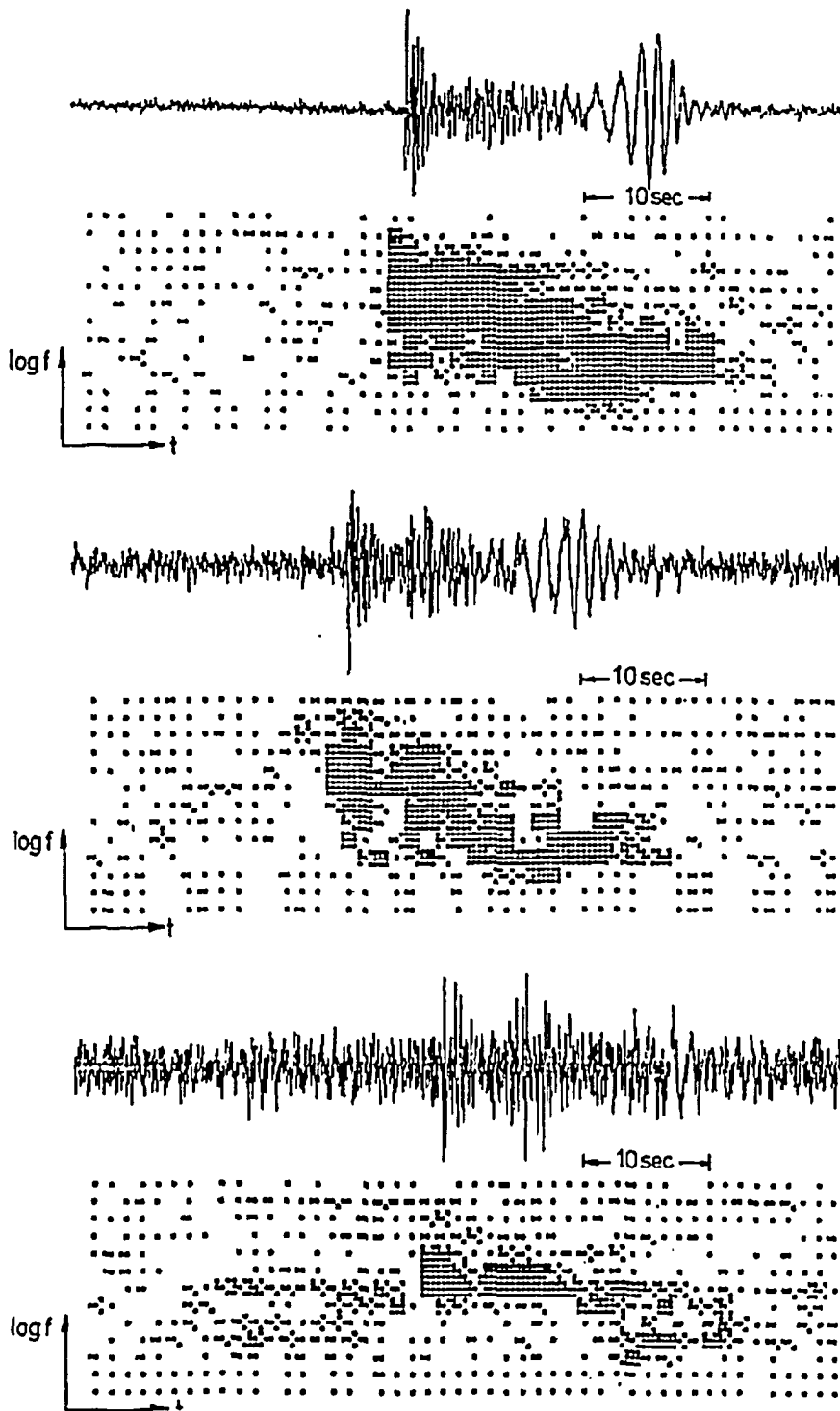
FIG. 4. All three examples are events of an area with 35 km epicentral distance. They differ significantly in their signal/noise ratio, and magnitude, but the principal structure of the related mental images stay stable. This invariance is a prerequisite for a recognition by a finite set of event patterns.

Dealing with black and white pictures reduces the pixel values to one bit, while the different scalings map all events of one type within a wider range of magnitude to one pattern.

The step of PR is then performed by crosscorrelation: For each threshold h_s the pattern fit between pattern M and picture matrix D_s at time lag τ is calculated by
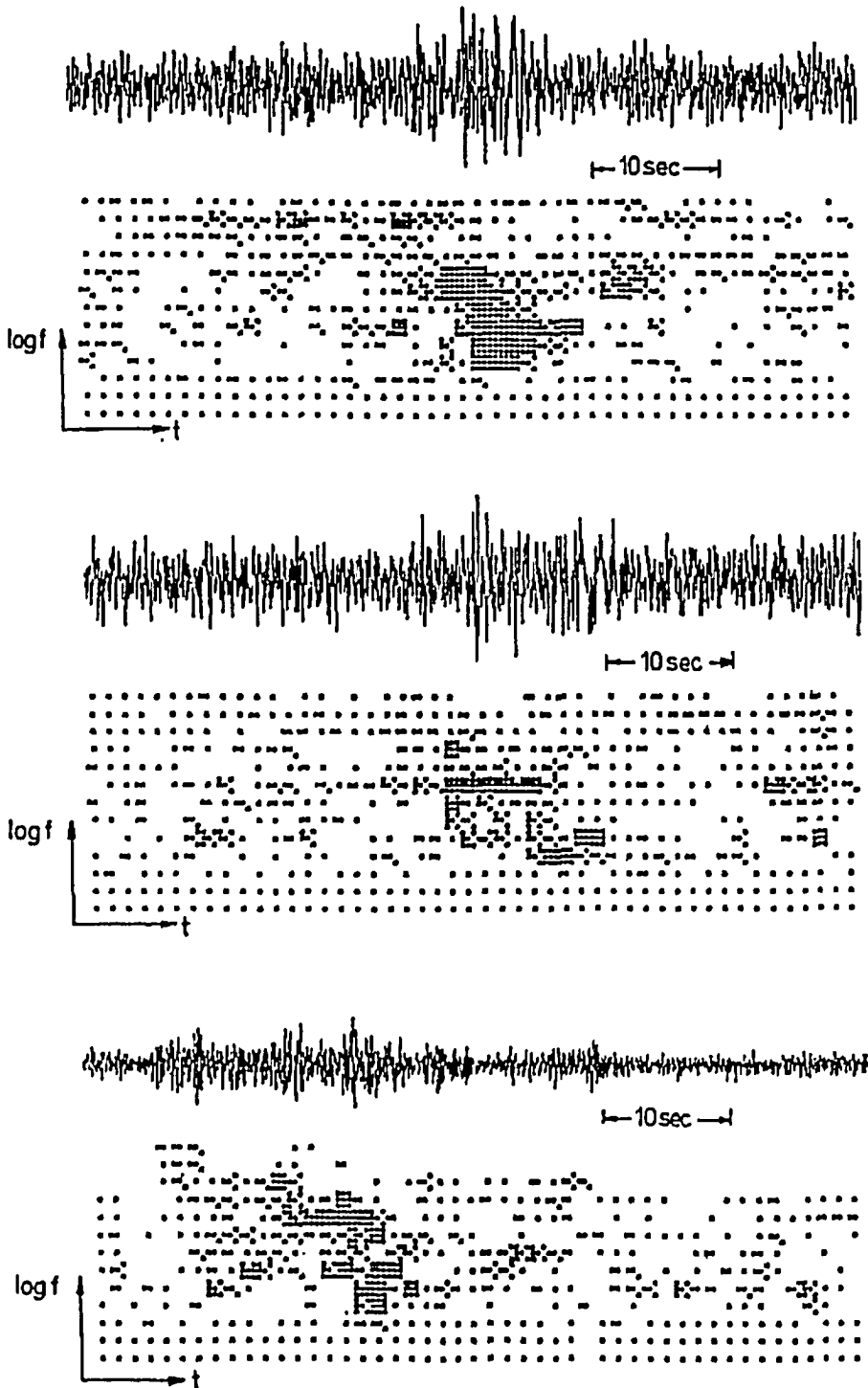
FIG. 5. Here the same situation as in Figure 4 is shown, but for epicentral distances of about 10 km and a much poorer signal/noise ratio. Remarkable is not the resulting detection, which will be performed by every detector with appropriate pre-whitening filter, but the clear contour of the mental images.

$$\text{fit}_s(\tau) = \frac{1}{N} \sum_i \sum_j m_{ij}\, d_{i,j+\tau}. \tag{12}$$

Note here that the multiplication can be replaced by fast look-up due to the discrete values of $(-1, 0, +1)$ for m_{ij} and d_{ij}. The factor $1/N$ from (8) normalizes fit_s to the interval $[-1, +1]$. A resulting $\text{fit}_s = +1$ demands energy above h_s for $m_{ij} = +1$ *and*
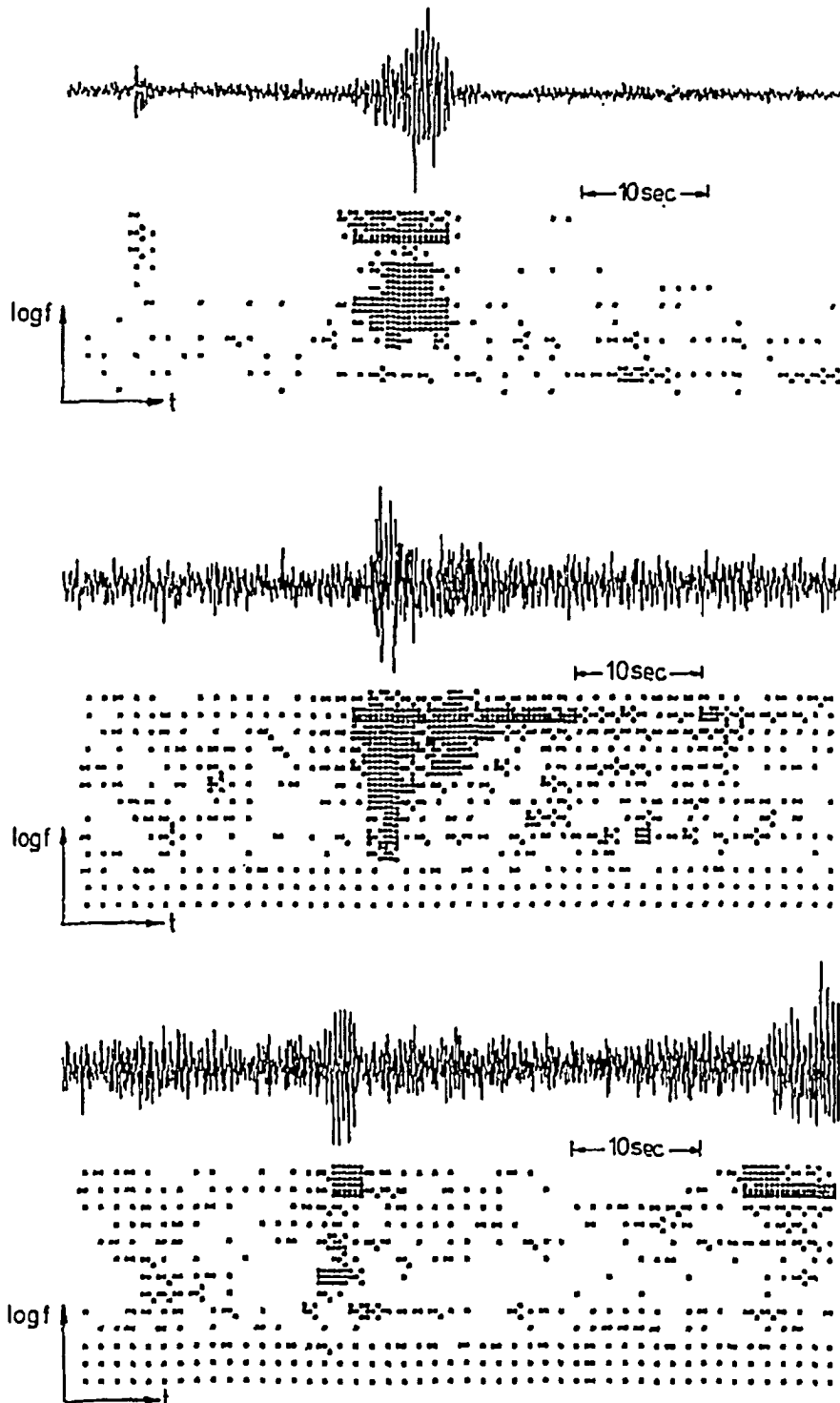
FIG. 6. Three noise bursts and the related mental images illustrate, how pattern recognition detectors can avoid noise detections independent of amplitude. All images show significant differences to the local events in Figure 5, although the seismograms of noise bursts will in general be most closely related to those of near local earthquakes.

no energy at $m_{ij} = -1$, while $\text{fit}_s = -1$ would mean inverse energy distribution. In case of noise and if balance condition (9) holds, all pictures with a random black and white distribution will yield $\text{fit}_s \cong 0$. So a common threshold for signal detection will be $\text{fit}_s \gtreqqless 0.5$.
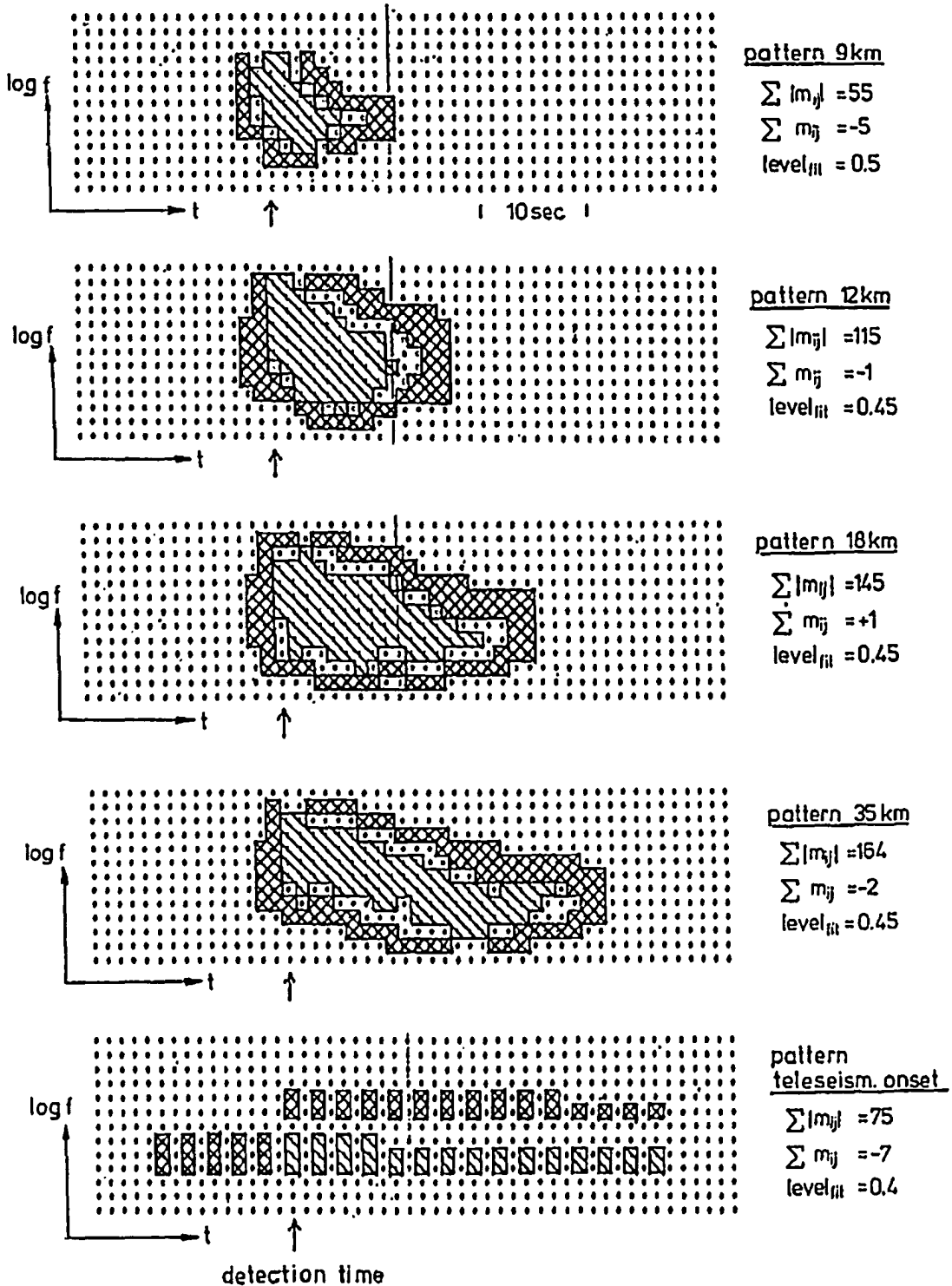
FIG. 7. The knowledge base of the pattern recognition detector consists of a set of distinct patterns. Each one represents a type of known seismograms from earthquakes or noise sources. The striped areas require signal energy above a common threshold, while in the crossed areas there must be *no* energy. The preliminary pattern set of the BUG-network shown here distinguishes events only by epicentral distance, not by different azimuths.

One last step has to follow PR and it is performed by a selection process: If an event occurs, raw detections are usually triggered within a wider range of time, for some thresholds h_s in (10) and for all similar patterns. So searching for the one maximum fit, ensures, that only one detection message per event is generated. This
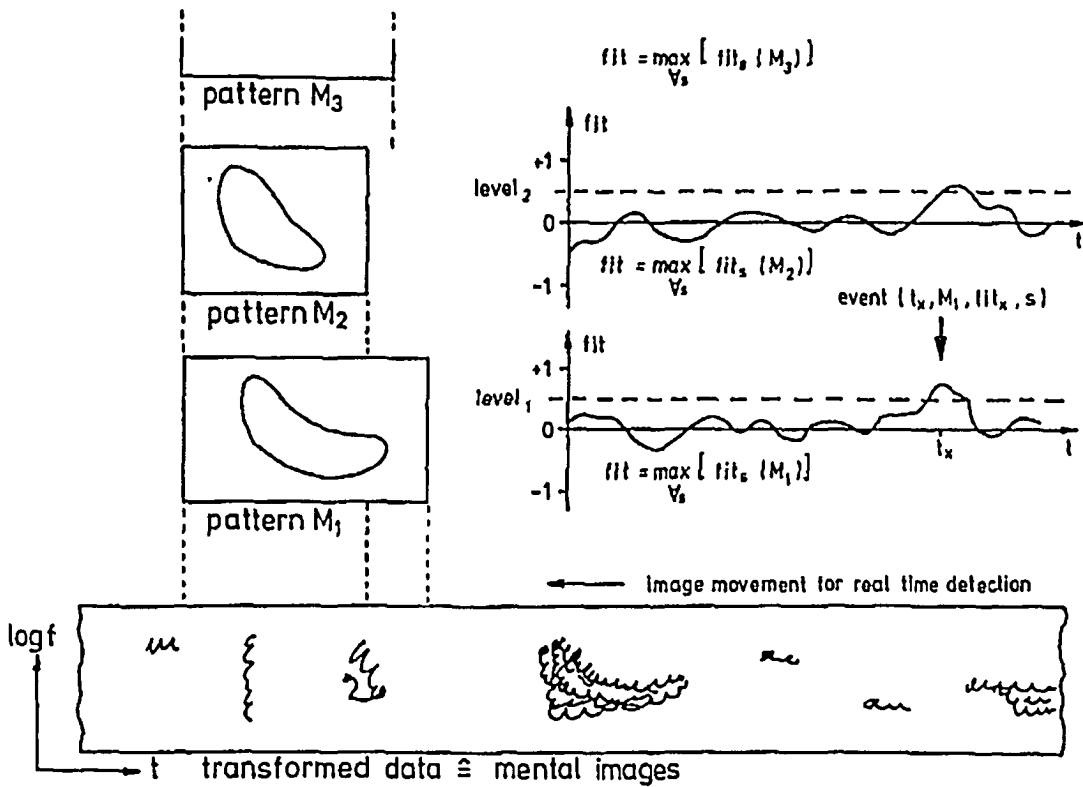
FIG. 8. The pattern recognition of earthquakes by mental images is like detection of visible objects on an assembly line. The mental image of the actual seismogram is shifted in time and compared to all predefined patterns. To trigger a detection, the fit between image and pattern must surpass a certain level. The final detection message is composed by a subsequent decision logic based on the maximum fit.

message includes detection time, pattern type, maximum pattern fit and threshold, so it supplies a subsequent coincidence detector or expert system with valuable information about amplitude and identification quality of the detected event.

Definition of the Similarity Measure

The similarity measure α_{xy} quantifies the correlation between two patterns X and Y. Except for the different normalization it represents the same relation between two matrices as the pattern fit in (12) and is defined by

$$\alpha_{xy} = \frac{1}{\sqrt{N_x N_y}} \max_\tau \left(\sum_i \sum_j x_{ij} y_{i,j+\tau} \right). \tag{13}$$

Calculation of the similarity measure can be used to avoid redundant pattern definitions. Table 1 shows α_{xy} for the pattern set of Figure 7. Values of 0.7 and more would indicate pattern proximity that slows down processing speed without raising the detection probability by additional information.

The similarity measure is also useful in a seismic network, when one event will not necessarily trigger the same type of pattern in all the neighboring stations. For a subsequent voting detector either all legal pattern combinations must be specified or calculating α_{xy} can check on reasonable similarity of the detected onsets.

TABLE 1

SIMILARITY α_{xy} OF BUG-PATTERNS

Typo	9 km	12 km	18 km	35 km	teleseis
9 km	1.00	0.30	0.26	0.15	0.05
12 km	0.30	1.00	0.58	0.39	0.10
18 km	0.26	0.58	1.00	0.53	0.13
35 km	0.15	0.39	0.53	1.00	0.21
teleseis	0.06	0.10	0.13	0.21	1.00

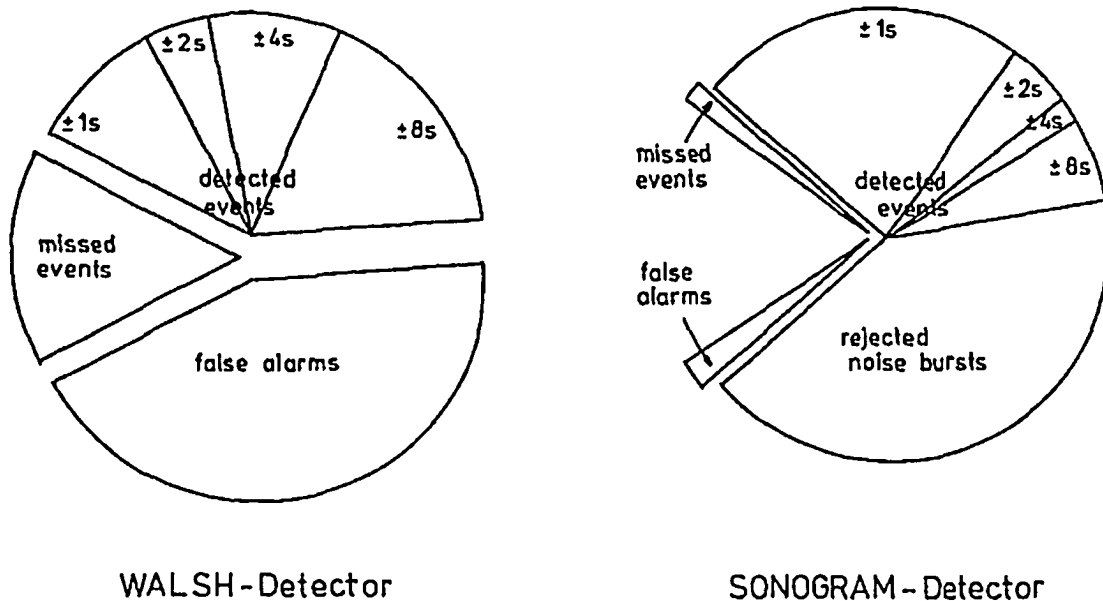WALSH-Detector   SONOGRAM-Detector

FIG. 9. The Walsh-detector had in 25 days on-line processing 125 false alarms, 118 detected earthquakes, and 44 missed events (half of them by network downtime). The Sonogram-detector improved timing accuracy for the earthquakes but rejected four very small ones. In addition it suppressed all of the noise bursts but six.

Test Installation

The first tests were performed off-line with 25 days of data from the BUG-network. The Sonogram-detector had to identify local seismicity in two minute time windows selected by the Walsh-detector. Number of missed events, false alarm rate and timing inaccuracies of the Walsh-algorithm for the 25 days are shown in Figure 9 and conform with other installations in high cultural noise. The resulting data set of 243 is representative for seismicity and noise bursts in the BUG-network except for some very small earthquakes that are missing.

The Sonogram-detector by means of the simple pattern set in Figure 7 identified all seismic events but four, rejected all noise bursts but six, and enhanced timing accuracy (Fig. 9). The PR approach is so sensitive that all earthquakes in Figures 4 and 5 are correctly identified and that the maximum amplitude of the four missed events is less than 6 dB above noise.

Although quite precise in timing, the Sonogram-detector is not a phase picker. Like all methods in frequency domain, it has an inherent uncertainty of half the window length. Instead it can be called a true event detector, since it utilizes the information of the whole seismogram to adjust exactly one detection time per event onto the P onset.

In contrast to the traditional Freiberger-based detector principles, PR-detectors

TABLE 2
BUG-PATTERNS VERSUS EPICENTRAL DISTANCE

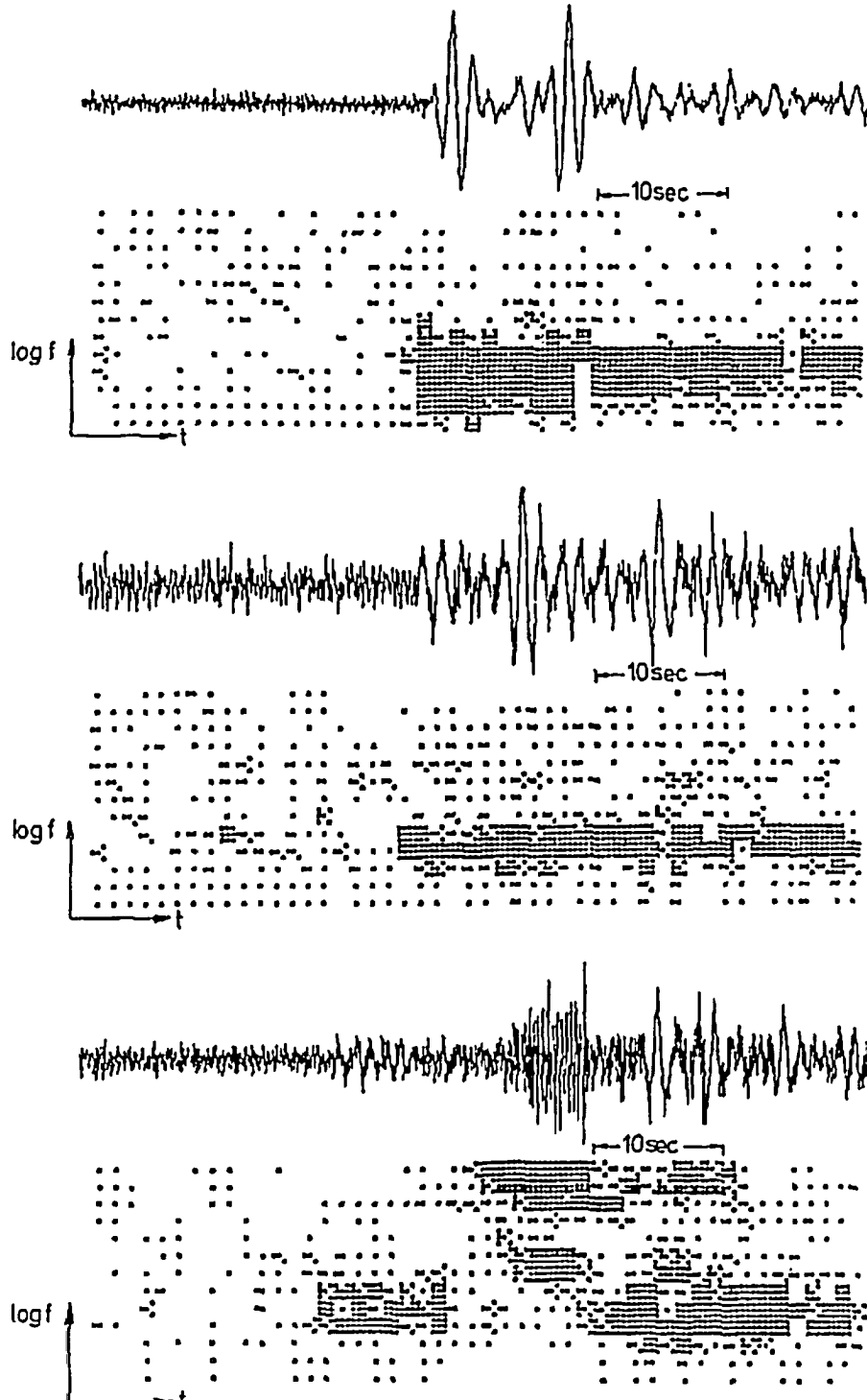| pattern distance	9 km	12 km	18 km	35 km	teleseis
0–10 km	1				
11–14 km	2	3	1		
15–25 km	2	3	3		
26–50 km		1	6	7	1
>1000 km					13

FIG. 10. Three examples of teleseismic P-wave onsets show the monochromatic spectrum for short period seismometer registrations. The advantage of PR in this application is the detection of noise as in the third image. A high frequent noise burst just appeared in the gap between P and pP of a Fiji earthquake ($\Delta = 153°$, $h = 64$ km, $m_b = 5.7$).

can be improved by updating the knowledge base rather than by changing the algorithms for computation or conclusion: so

- Definition of additional noise patterns allows a lower detection threshold without false alarm increase (only the most similar pattern is picked),
- Refinement of the event patterns gives better identification for weak events. An analysis of the detected events by identified pattern type versus determined epicentral distance still shows significant uncertainty (Table 2). The inaccuracies in onset time (see Fig. 9) and distance association are due to the azimuthal distribution of earthquakes which was neglected in the initial set of purely distance-dependent patterns in Figure 7.

Thus with PR, the detector performance depends crucially on the definition of an appropriate knowledge base. Starting with a rough estimate during installation, a successful detector refinement promises performance that equals the results of a skilled human observer.

Detection of Teleseisms

Like the detection of local events, detection of teleseismic earthquakes in general can be accomplished by defining patterns on the whole event. This will scale the FFT window to the frequency range of broadband seismometers.

On short-period seismogram traces, however, the power of PR is quite useless, since shifting of spectral energy cannot be observed. The P onset appears as a monochromatic sweep, while S- and surface-waves are outside the frequency range (Fig. 10). To detect teleseismic events in this situation, the lowest pattern of Figure 7 essentially copies a STA/LTA-based detector: It demands energy above the threshold within a constant frequency range (i.e., bandpass-filtering) and for a certain duration (i.e., postdetection by time window). The effect of edge conversion—no simultaneous high frequency energy—is that of a second STA/LTA acting on high frequencies with a negative trigger weight (Evans and Allen, 1983).

CONCLUSIONS

PR-detectors form a new class of processing schemes based on positive decision logic and are not related to the traditional Freiberger-based detectors.

The Sonogram-detector uses images of spectral energy versus time to recognize known patterns of earthquakes. It has high detection probability and timing accuracy and, most notably, the ability to reject known noise bursts.

The principal restriction for all PR-methods is that only those signals known *a priori* can be detected. So PR-detectors will not supersede the traditional detectors with their negative decision logic under all circumstances. The advantage of the latter is their use "off the shelf," the reliable detection of all temporary changes without any additional knowledge. These properties make them ideal for every short-term field application.

The consequence of the trade-offs mentioned above is to combine the power of both approaches in a universal detector scheme. This can be done in a pre/postdetector design as for the test installation, where sonograms are calculated on the preselected time windows. Also possible is a combination in parallel, where the sonograms are updated continuously. Since the PR-detector acts as a matched filter, it should have a slightly higher sensitivity for known signals than the traditional detectors.

The PR-detection on mental images is a general concept not at all limited to sonograms. For example vespagrams are powerful analytical tools for seismic arrays displaying energy in time versus slowness for fixed azimuths (Davies et al., 1971). They are well suited for PR, and Capon and Davies (1971) even proposed this idea for detection purposes. They did not apply image processing techniques, however, and the approach was forgotten. Today, a Vespagram-detector can be based on processing steps of the Sonogram-detector and is another promising application of PR on mental images.

ACKNOWLEDGMENTS

REFERENCES

Allen, R. (1978). Automatic earthquake recognition and timing from single traces, Bull. Seism. Soc. Am. 68, 1521-1532.

Allen, R. (1982). Automatic phase pickers: their present use and future prospects, Bull. Seism. Soc. Am. 72, S225-S242.

Anderson, K. R. (1981). Glitch Noise Recognition using Augmented Transiton Network Grammars, Semiannual Technical Summary Report, pp. 19-21, MIT Lincoln Lab., Lexington, Mass., April-Sept.

Anderson, K. R. (1982a). Syntactic analysis of seismic waveforms using augmented transition network grammars, Geoexploration 20, 161-182.

Anderson, K. R. (1982b). Syntactic Analysis of Regional Seismic Signals, Semiannual Technical Summary Report, pp. 27-30, MIT Lincoln Lab., Lexington, Mass.

Blandford, R. R. (1982). Seismic event discrimination, Bull. Seism. Soc. Am. 72, S69-S87.

Capon, J. and D. Davies (1971). Pattern Recognition Techniques for the Interpretation of Vespagrams and Beamsplit Data, Semiannual Technical Summary Report, pp. 38-41, MIT Lincoln Lab., Lexington, Mass.

Chen, C. H. (1984). On a segmentation algorithm for seismic signal analysis, Geoexploration 23, 35-40.

Davies, D., E. J. Kelly, and J. R. Filson (1971). The Vespa process for the analysis of seismic signals, Nature 232, 8-13.

Evans, J. R. and S. S. Allen (1983). A teleseismic-specific detection algorithm for single short-period traces, Bull. Seism. Soc. Am. 73, 1173-1186.

Faure, C., H. Soldano, and Th. van der Pyl (1984). Descriptive methods and processing of seismic signals, Geoexploration 23, 17-34.

Fleck, P. L., Jr., and L. T. Fleck (1964). Studies using Sonograms, Semiannual Technical Summary Report, pp. 11-14, MIT Lincoln Lab., Lexington, Mass.

Freiberger, W. F. (1963). An approximate method in signal detection, Quarterly Appl. Math. 20, 373-378.

Gaby, J. E. and K. R. Anderson (1984). Hierarchical segmentation of seismic waveforms using affinity, Geoexploration 23, 1-16.

Goforth, T. and E. Herrin (1981). An automatic seismic signal detection algorithm based on the Walsh transform, Bull. Seism. Soc. Am. 71, pp. 1351-1360.

Gonzales, R. C. and P. Wintz (1987). Digital Image Processing, Addison Wesley, Reading, Massachusetts.

Haugeland, J. (1985). Artificial Intelligence: The Very Idea, MIT Press, Cambridge Massachusetts.

Joswig, M. (1987). Methoden zur automatischen Erfassung und Auswertung von Erdbeben in seismischen Netzen und ihre Realisierung beim Aufbau des lokalen Bochum University Germany, Netzes, Dissertation, Inst. f. Geophysik, Ruhr-Univ., Bochum.

Joswig, M. and H.-P. Harjes (1986). A seismic network with loosely coupled remote stations, in Proc. XX. General Assembly of ESC 1986 in Kiel, D. Mayer-Rosa and J. M. Van Gils (Editors), H. Stiller, Zurich.

Kennett, B. L. N. (1975). Short-term spectral analysis and sequency filtering of seismic data, in Exploitation of Seismograph Networks, K. G. Beauchamp (Editor), Noordhoff, Leiden, pp. 283-296.

Kosslyn, S. M. (1980). *Image and Mind*, Harvard Univ. Press, Cambridge, Massachusetts.

Liu, H.-H. and K.-S. Fu (1983). An application of syntactic pattern recognition to seismic discrimination, *IEEE Trans. Geosc. Rem. Sens.* GE-21 (2), 125–132.

Murdock, J. N. and C. R. Hutt (1983). *A New Event Detector Designed for the Seismic Research Observatories, U.S. Geol. Surv. Open-File Rept. 83-785.*

Sheppard, R. M. (1978). Sonograms of SRO Broad-Band Recorded Events, Semiannual Technical Summary Report, MIT Lincoln Lab., Lexington, Massachusetts, pp. 9, 20–31, March.

Stewart, S. S. (1977). Real-time detection and location of local seismic events in central California, *Bull. Seism. Soc. Am.* 67, 433–452.

Swindell, W. H. and N. S. Snell (1977). Station Processor Automatic Signal Detection System, Phase I: Final Report, Station Processor Software Development, Report ALEX(01)-FR-77-01, AFTAC Contract F08606-76-C-0025, Texas Instruments, Dallas.

Vanderkulk, W., F. Rosen, and S. Lorenz (1965). Large Aperture Seismic Array Signal Processing Study, IBM Final Report, ARPA Contract SD-296, Rockville, Maryland.

Winston, P. H. (1984). *Artificial Intelligence*, Addison-Wesley, Reading, Massachusetts.

INSTITUT FÜR GEOPHYSIK
RUHR-UNIVERSITÄT BOCHUM
POSTFACH 102148
4630 BOCHUM 1, F.R. GERMANY