EARTHQUAKE/NOISE DISCRIMINATION AND ESTIMATION OF P-S PHASES BASED ON WAVE CHARACTERISTICS

Kokinou E.[1], Panagiotakis C.[2], and Vallianatos F.[1]

[1] *Laboratory of Geophysics and Seismology, Department of Natural Resources and Environment, Technological Educational Institute Crete, 3 Romanou Str. Chalepa, Chania, Crete, GR 73133 Greece, Email: ekokinou@chania.teicrete.gr*

[2] *Computer Science Department, University Of Crete, P.O. Box 2208, Heraklion,, Greece, Email: cpanag@csd.uoc.gr*

Abstract

Aim of the present work is to adapt and develop an algorithm for extracting first motion information from earthquake seismograms. In specifics, the present paper deals with the characterization of the z component of a given seismogram in order to discriminate an earthquake signal from noise recording.

Segmentation and further classification of the seismogram comprises the initial step. The earthquake detection (P arrival) is performed using a low computation cost method based on signal energy. Furthermore, accurate picking of S arrival is based on signal energy and central frequency estimation.

Experimental results show satisfied efficiency of the pre-mentioned scheme, low processing time and potential application in real-time environment.

Key words: *Automatic picking, Seismogram, Signal segmentation.*

Περίληψη

Σκοπός της παρούσας εργασίας είναι η ανάπτυξη αλγόριθμου, με στόχο να εξάγει πληροφορία από σεισμολογικές καταγραφές. Πιο συγκεκριμένα χρησιμοποιείται η κατακόρυφη συνιστώσα για να διαχωριστεί το σήμα που αντιπροσωπεύει ένα σεισμό από τον θόρυβο.

Πρώτο βήμα, στην διαδικασία που παρουσιάζεται, είναι η τμηματοποίηση και ταξινόμηση του σήματος. Ακολουθεί ο εντοπισμός των πρώτων αφίξεων των P και S κυμάτων με βάση τα χαρακτηριστικά του σήματος, όπως η ενέργεια και η συχνότητα. Ο εντοπισμός της εισόδου του P κύματος απαιτεί γενικά μικρό υπολογιστικό κόστος.

Τα πειραματικά αποτελέσματα της εφαρμογής του παραπάνω αλγόριθμου αποδεικνύονται ικανοποιητικά, ο υπολογιστικός χρόνος είναι μικρός και η εφαρμογή του συγκεκριμένου αλγόριθμου σε πραγματικό χρόνο εφικτή.

Λέξεις κλειδιά: Αυτόματος εντοπισμός, Σεισμόγραμμα, Τμηματοποίηση σήματος.

1. Introduction

Accurate picking of P and S phases comprises the most important step for earthquake location and tomographic study (Kokinou *et al.* 2006) and for any further understanding of crustal and upper

mantle structure. Large data sets must be analysed in order to reveal P arrival times and further construct the seismic tomography for the studied area.

According to Di Stefano *et al.* (2006) two main approaches are basically used for automatic phase picking. The first one comprises the single-trace picking, i.e. each seismogram is picked independently from the other. The waveform attributes (amplitude, frequency, polarization) are used for the seismogram characterization. The well known STA/LTA (short-term-average long-term-average) algorithm (Allen 1978, 1982) applied on an envelope function sensitive to both amplitude and frequency of seismograms still comprises part of Earthworm (Johnson *et al.* 1994) and Sac2000 (Goldstein *et al.* 1999).

The second approach comprises a multitrace method, i.e. several seismograms at once based on the waveform similarity (Aster and Rowe 2000) but it still needs further development in adaptive filtering and event clustering (Rowe *et al.* 2002).

In the context of the present work we have developed an automatic P-S phase picker, sensitive in small period and amplitude changes, applied in short period seismograms. The general concept for developing the automatic picker is based on an algorithm applied successfully in sound analysis (Panagiotakis and Tziritas 2005). The signal energy and the central frequency comprise the basic wave characteristics to be analysed in order to be used in the discriminator. Thus, we use the squared signal amplitude, equivalent to the signal energy and the central frequency, estimated by the local minima and maxima detector (Panagiotakis *et al.* 2006).

The 8th January 2006 earthquake (mb=6.7, CMT Harvard routine analysis) and its aftershocks, occurred in northwestern offshore area of Crete Island were recorded by the stations of the regional seismological network (Fig. 1) of the laboratory of Geophysics and Seismology of the Technological and Educational Institute of Crete. The vertical component z of the pre-mentioned earthquake event is used as an example presented in next paragraphs. Lower magnitude earthquake seismograms, as well as pure noise seismograms are also used in order to prove the proposed algorithm stability and robustness.

The rest of the paper is organized as follows: Section 2 presents the methodology. Experimental results and discussion are provided in Sections 3 and 4, respectively.
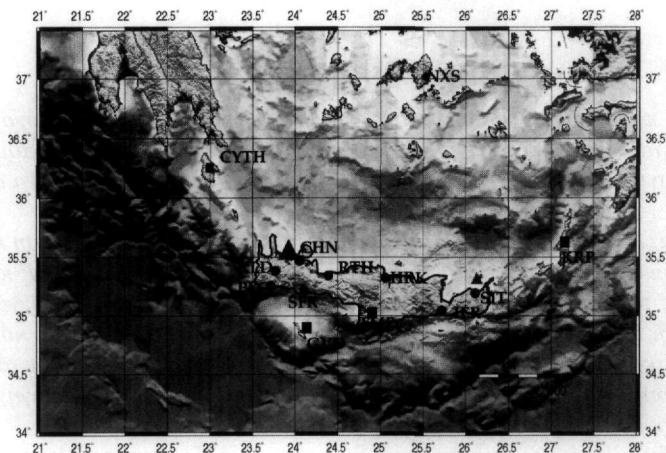
Figure 1 - Geographical distribution of the Seismological Network of south Aegean

2. Methodology

2.1. Algorithm overview

The proposed algorithm consists of several main modules (Fig. 2). As an input, the z component of the seismogram (z signal) for a time period (e.g. 15 minutes) is used.

First, P-picking algorithm is executed yielding the P arrival time of the recorded earthquake based on signal energy analysis. Further the S arrival time is computed using a frequency-energy based method.

At each module, the "quality" of the estimation is computed yielding a value concerning the reliability of the estimation. Therefore, if the given signal does not contain any event (is just noise), then P-picking module gives low quality of the P arrival time.
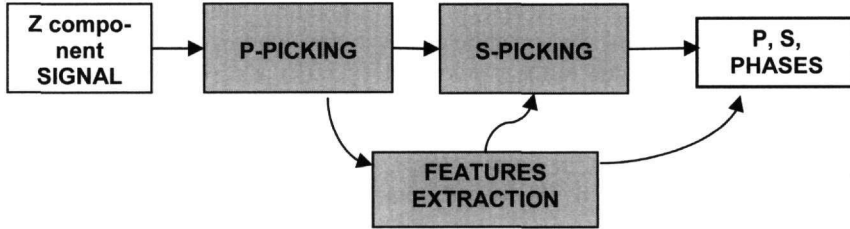
Figure 2 - Scheme of the proposed system architecture

2.2. Picking of P arrival

The P arrival time of the recorded earthquake is characterized by the property that the energy of a time window locating prior P arrival time is lower than the energy of a time window locating after P arrival time. This property has been used in our method. Let f(t) be the ratio between the signal energy of a time window locating after P arrival time t and the signal energy of a time window locating prior P arrival time t. Equation (1) describes the function f(t).

Equation 1 – Formula for f(t) definition

$$f(t) = \frac{\sum\limits_{k=t}^{t+W-1} z^2(k)}{\sum\limits_{k=t-W}^{t-1} z^2(k)}$$

Both windows are of equal size (W) and in the present work have been selected to be less than 5 seconds. The global maximum of f(t) corresponds to the most possible P arrival time while the value f(P) gives the "quality" of the estimation.

2.3. Picking of S arrival

The S arrival time, located after P, is characterized by the property that the central frequency estimated on a time window locating prior S arrival time is higher than the central frequency estimated on a time window locating after S arrival time. At the same time, the signal energy changes prior and after S arrival. In order to estimate the central frequency or period in a time window of the z component of the recorded signal, the signal local maxima/minima are estimated. Therefore, the mean period is evaluated by getting the mean difference in time between maxima and/or minima.

Let $T_1(t)$ and $T_2(t)$ be the two estimated periods by the local minima-maxima proposed method of a time window locating prior t and after t, respectively. Let g(t) be the function defined as follows,

Equation 2 – Formula of g(t) definition

$$g(t) = (T_1(t) - T_2(t))^3 \bullet \left| \sum_{k=t}^{t+W'-1} z^2(k) - \sum_{k=t-W'}^{t-1} z^2(k) \right|, \; t \geq P$$

The third power of $T_1(t)$-$T_2(t)$ has been used, since as our experimental results show, it yields better results. Thus, the increase of signal period after S arrival is a robust feature, comprising the most important factor in g(t) definition.

The window size (W') has been selected to be the duration which corresponds to four local maxima of z. g(t) measures how the signal belonging to the window prior t {$z(k)$, $k \in [t\text{-}W', t\text{-}1]$} differs from the signal belonging to the window after t {$z(k)$, $k \in [t, t+W'\text{-}1]$}, using period and energy features. Thus, the window size (W') dynamically changes in respect to time. According to the S properties, the first maxima of g(t) that is greater than a adaptive threshold value (e.g. 15% of global maximum) corresponds to the S arrival time, while the value of g(S) gives the "quality" of the estimation. In order to speed up the computation, g(t) is estimated on local maxima and minima of z.

2.4. Method advantages and implementation

The method is based on robust features and it is simple and fast. We use just energy and local maxima – minima of the signal in order to accurate estimate the central frequency, since the given signal is normally noisy, so more complicated features will be affected by the very low signal to noise ratio.

The proposed scheme yields the "quality" of estimation. This value corresponds to the reliability of estimation and it can be used to recognize if the given signal corresponds to an earthquake signal or not. Moreover, the whole procedure is mainly automatic. It does not require any parameter. However, the user can change easily the window sizes, or modify the output file.

The method has been implemented using Matlab. We have used a module based implementation as it is illustrated in Figure 2. For our experiments, we have been using a Pentium 4 CPU at 2.8 GHz. A typical processing time for the execution of the proposed scheme is about 2 minutes to analyze an hour signal. The most of the processing time (80-90 %) is consuming on P picking. However, this procedure can be changed using a hierarchical estimation of two stages (Panagiotakis and Tziritas 2005) in order to speed up the method.

3. Results

3.1. Evaluation of the results

The pre-mentioned scheme has been tested extensively using as input the z component of seismograms, recorded after the Kythira earthquake (8 January 2006). Examples of them are illustrated in the following figures. The P and S arrivals are marked in Figures 3 to 7 by a diamond and a square symbol, respectively. In addition, the f(t), g(t) functions of each seismogram are illustrated.

Figure 3 shows the earthquake of 8th January 2006 (at 11:35:09.89 UTC) of mb=6.7 according to CMT Harvard routine analysis. The values f(P), g(S) (Fig. 3c, 3d) of the specific event, are 4400 and $7.5 \cdot 10^{17}$ respectively, quite high in comparison to the corresponding values of f(P) and g(S) for the aftershocks. Similarly, Figures 4 to 7 show four aftershocks of the pre-mentioned event occurred in the next 6 hours after the main event. Noise examples are presented in Figures 8 and 9.
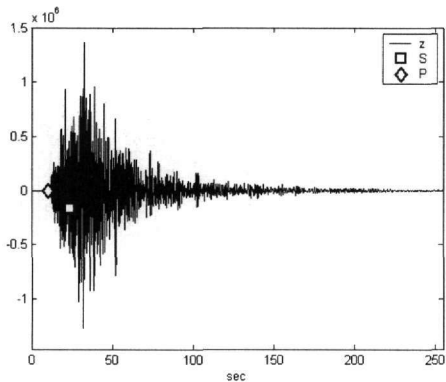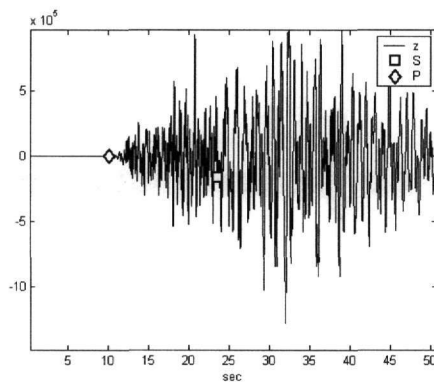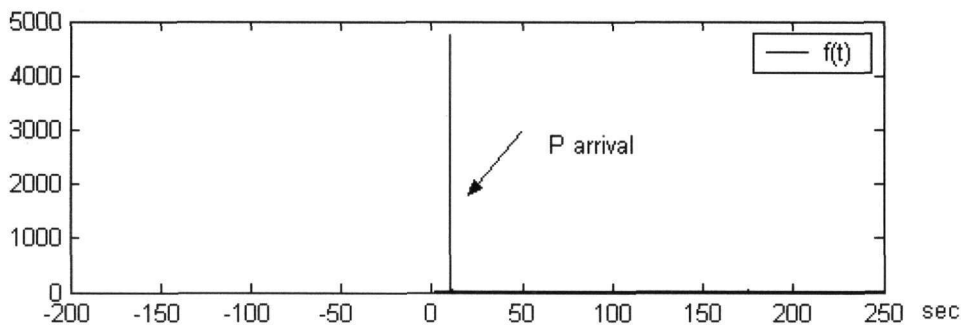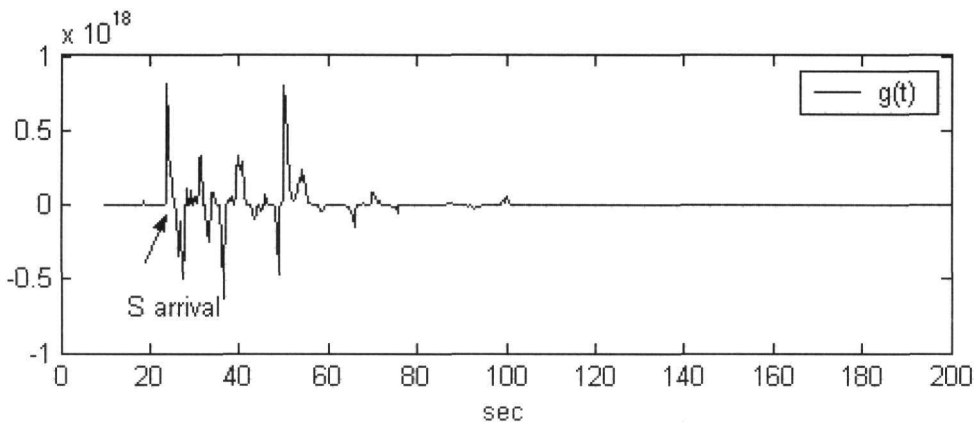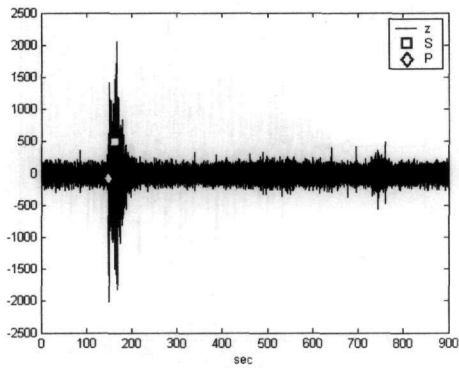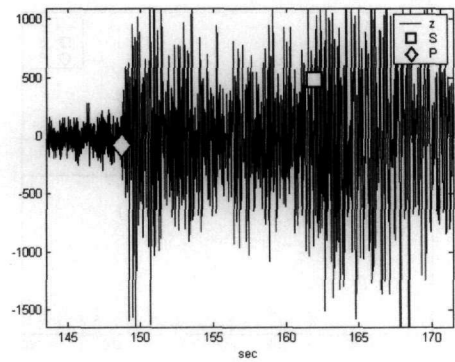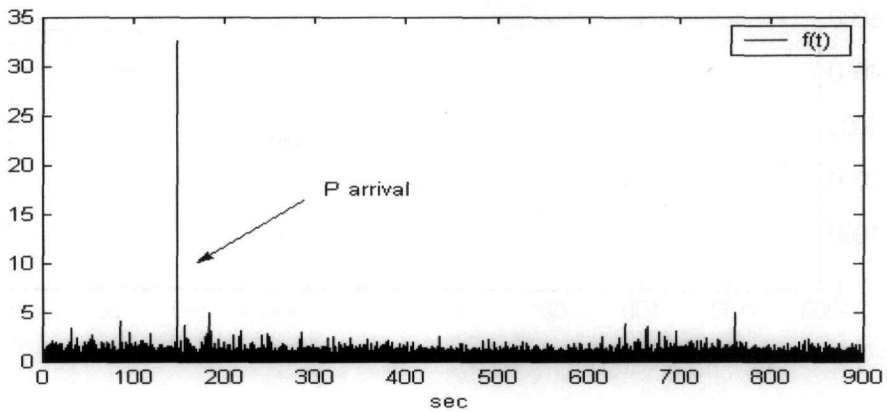
Figure 3 - (a) P and S arrivals are projected on z component of the 8th January 2006 (11:35:09.895 UTC, mb=6.7, CMT Harvard routine analysis) earthquake signal, as estimated by the proposed scheme, (b) enlarged image in order to demonstrate the picking precision of the proposed scheme, (c) the f(t) function, (d) the g(t) function
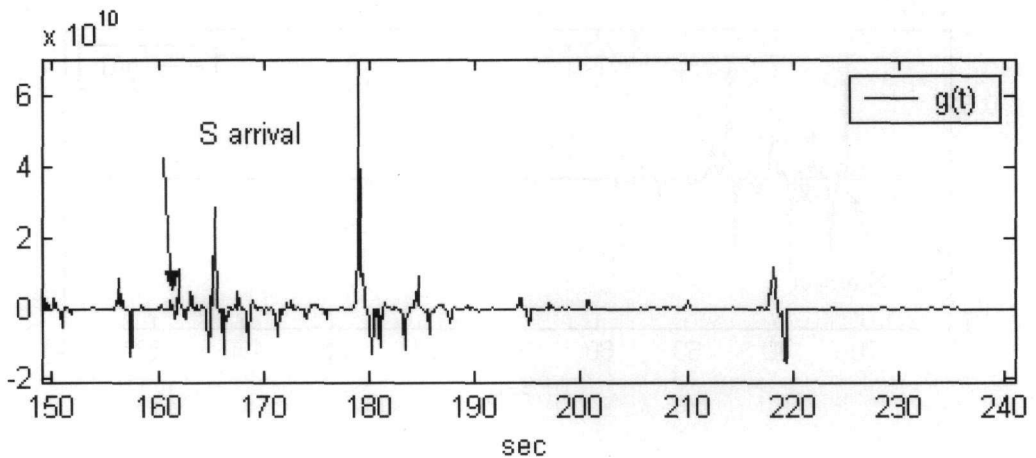
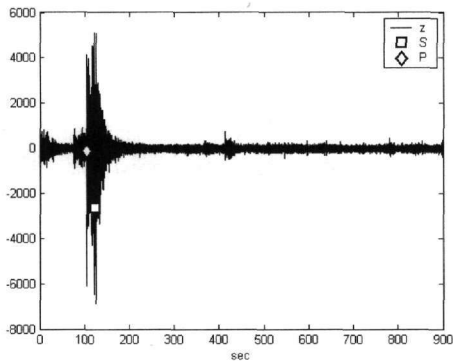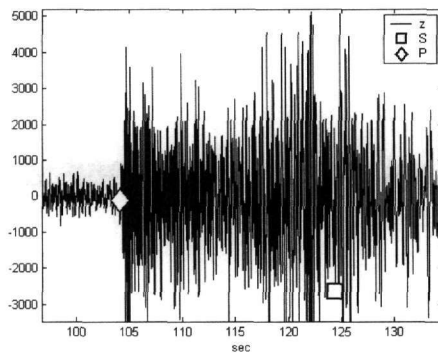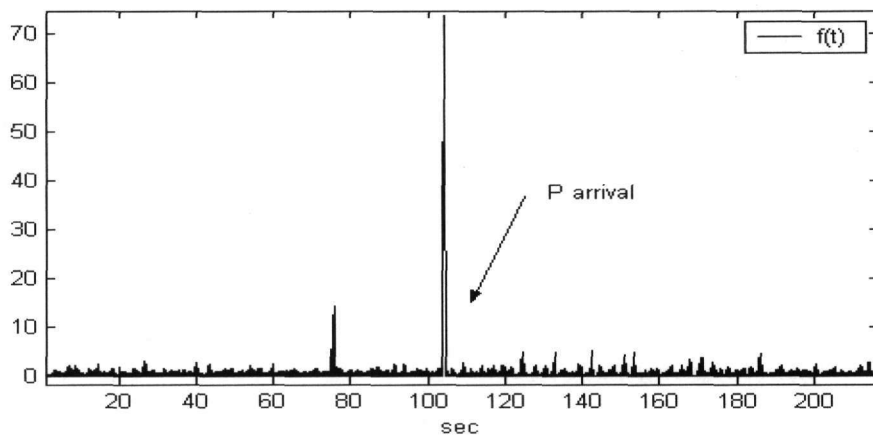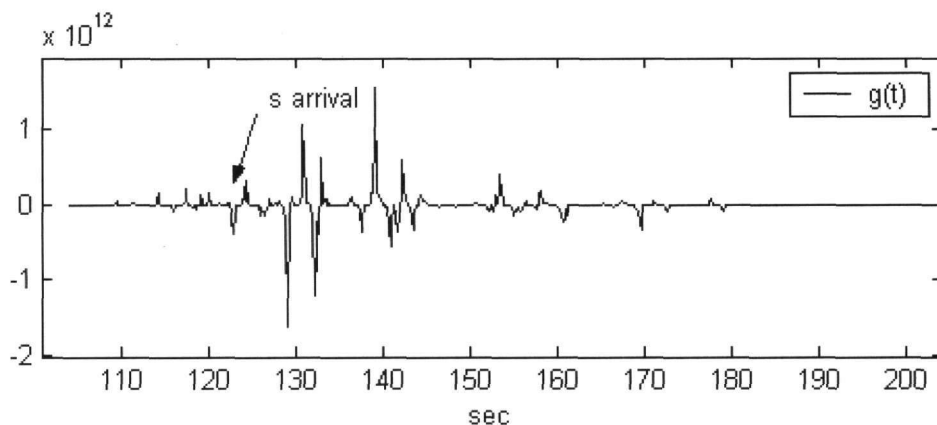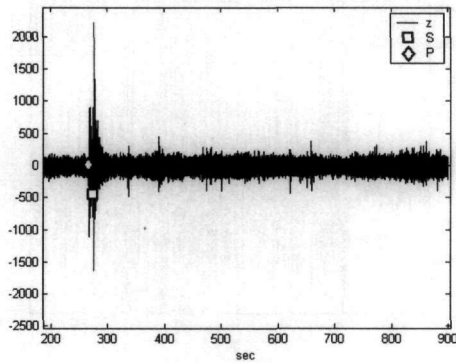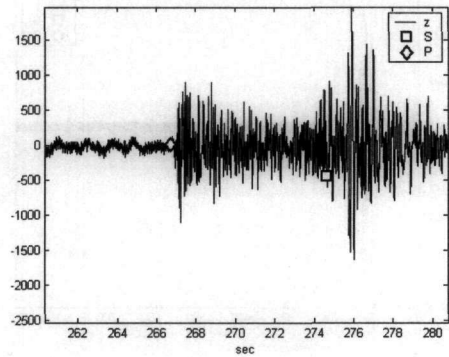Figure 4 - (a) P and S arrivals are projected on z component of the 8th January 2006 (11:52:28.503 UTC, M_L = 3.3 NOA) earthquake signal, as estimated by the proposed scheme, (b) enlarged image in order to demonstrate the picking precision of the proposed scheme, (c) the f(t) function, (d) the g(t) function
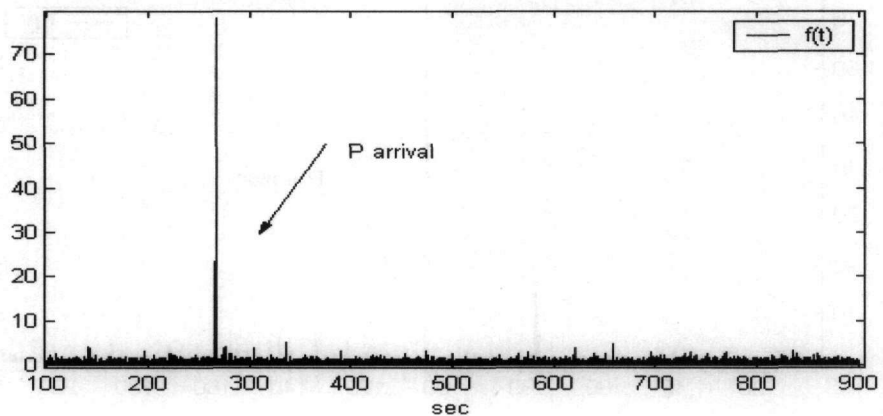
(a)

(b)

(c)

(d)

Figure 5 - (a) P and S arrivals are projected on z component of the 8th January 2006 (12:06:43.935 M$_L$ = 3.5 NOA) earthquake signal, as estimated by the proposed scheme, (b) enlarged image in order to demonstrate the picking precision of the proposed scheme, (c) the f(t) function, (d) the g(t) function
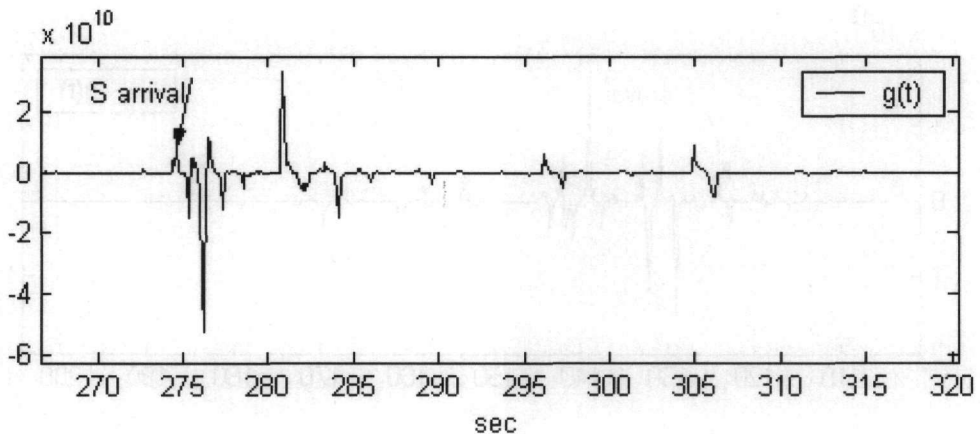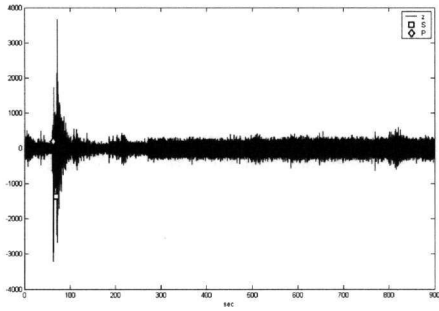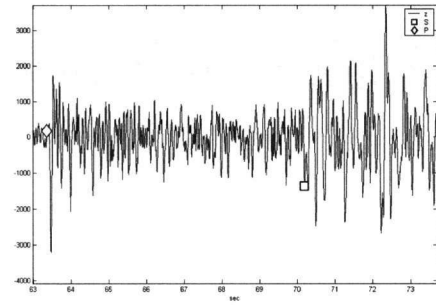
Figure 6 - (a) P and S arrivals are projected on z component of the 8th January 2006
(12:24:26.519, not found in NOA and EMSC catalogs) earthquake signal, as estimated by the
proposed scheme, (b) enlarged image in order to demonstrate the picking precision of the
proposed scheme, (c) the f(t) function , (d) the g(t) function

(a)

(b)

(c)

(d)

Figure 7 - (a) P and S arrivals are projected on z component of the 8th January 2006 (13:06:03.167, M$_L$ = 3.1 NOA) earthquake signal, as estimated by the proposed scheme, (b) enlarged image in order to demonstrate the picking precision of the proposed scheme, (c) the f(t) function, (d) the g(t) function
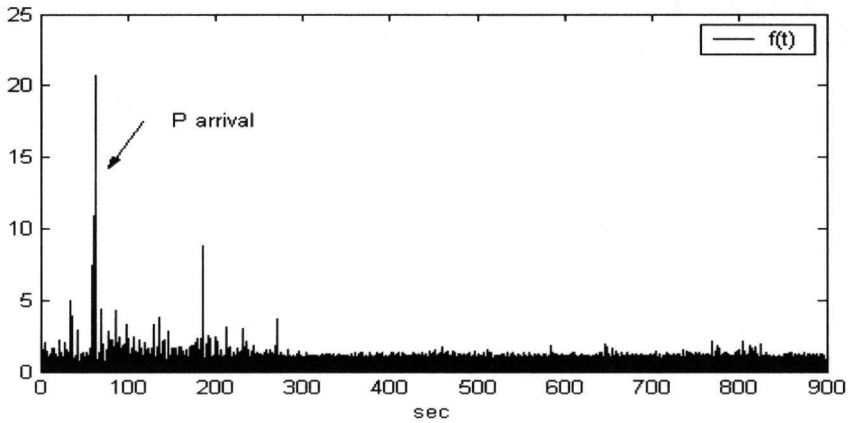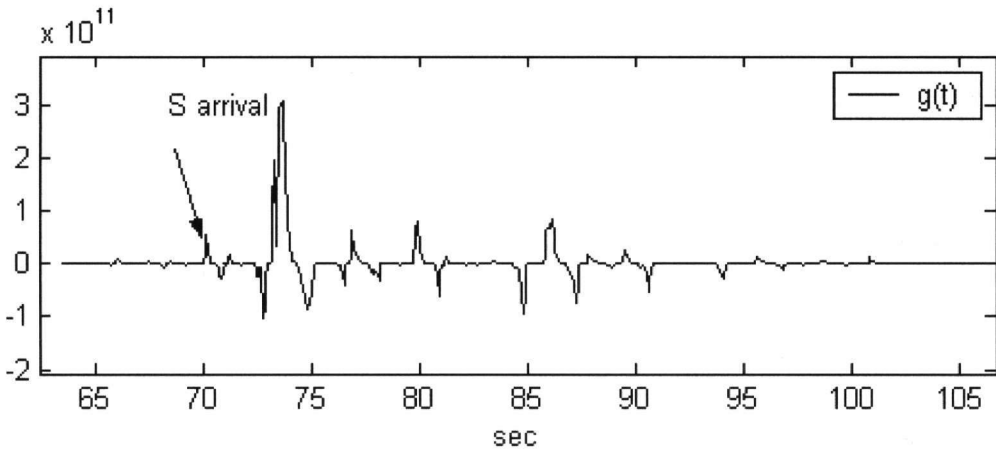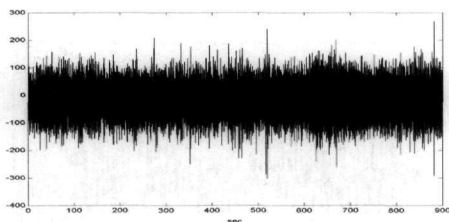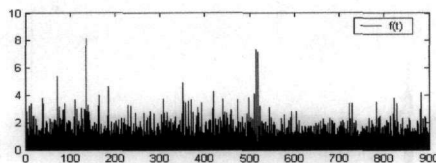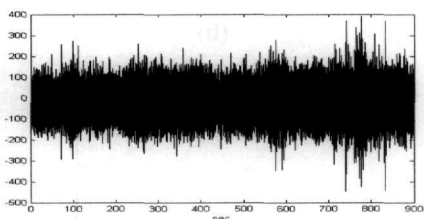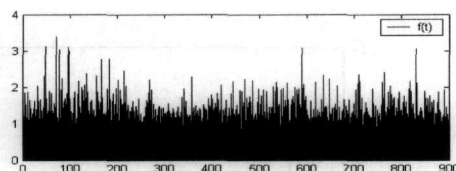
Figure 8 - (a) noise recording on z component, (b) the f(t) function of the previous noise re-cording

Figure 9 - (a) an other example of noise recording on z component, (b) the f(t) function of the previous noise recording

The function f(t), for these examples, gets values less than 10 (8 and 4, respectively) quite low in comparison to the corresponding values of f(P) of the earthquake signals. So, a threshold value of about 10 could be used in order to automatically reject the noise signals.

The detection probability for an earthquake (P arrival detection), independently from the accuracy of P estimation, at the present phase of the proposed system development is almost 100% in cases of strong to moderate events and reduces gradually in cases of weak and low energy events. This decision refers on a negative or positive answer concerning the given seismogram and is extracted by a threshold on the maximum f(t) value.

In order to evaluate the P and S pickings, a database of 40 earthquake recordings from 6 stations (Chania, Rethymno, Heraklion, Sfakia, Ierapetra and Sitia, Fig. 1) located on Crete Island, occurred in the time period between 8th February 2006 and the end of April 2006 in the offshore region around Crete, was created. The percentages of P and S pickings, which differences with the manual ones are below a given threshold [e.g. 0.1 second (Wang and Teng 1997)], are 92.5 % and 73 %, respectively. The wrong P or S detection was due to low energy seismograms, because we decided to use very noisy data in order to find out the limitations of the method. Additionally, even in cases of false P or S detection it was observed that the picks belonged to the part of the earthquake signal.

Concerning the accuracy of the P and S arrivals detection we have also used conventional software (PQL seismic trace viewer application) for seismogram analysis and the results are almost the same with those of the proposed method. There is also the possibility for the S arrival time to be picked on the N-S or E-W component.

In Figure 10 is illustrated the P and S detection error (in seconds) in respect to noise to signal ratio for the 8th January 2006 (09.895 UTC) earthquake z signal. It is observed that the error smoothly increases with respect to the noise energy. It holds that the maximum error of P and S is two and one seconds for 50% noise to signal ratio, respectively.
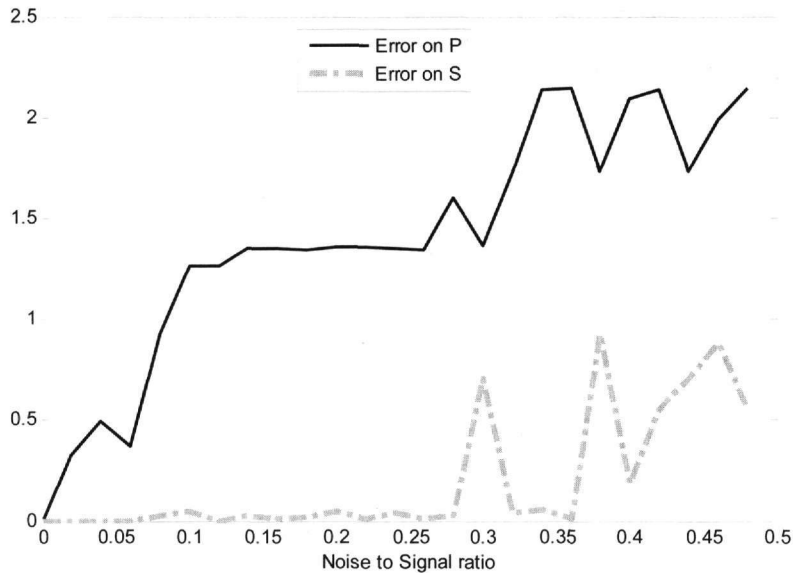
Figure 10 - P and S detection error (in seconds) in respect to noise to signal ratio for the 8th January 2006 (11:35:09.895 UTC) earthquake z signal

4. Conclusions

In this paper we adapt and develop a fast and effective algorithm, applied in sound analysis to discriminate speech and music, in order to detect first P and S arrivals in seismograms. The proposed algorithm is based on signal energy and frequency. P arrival is detected from signal energy while S arrival detection is implemented using the energy and central frequency, estimated on a time window. The detection of both P and S arrivals is controlled by "quality" factors [f(t) and g(t) functions], concerning the reliability of the estimation.

The main contribution of this work concerns the proposed seismogram analysis using robust and simple energy-frequency based features that suffice for an earthquake detection and high time accuracy of P and S arrival estimation.

As future work, we plan to extend the proposed method on P converted phases estimation between the P and S first arrivals.

5. Acknowledgments

6. References

Allen, R., 1978. Automatic earthquake recognition and timing from simple 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.

Aster, R., and Rowe, C., 2000. Automatic phase pick refinement and similar event association in large seismic datasets, In C.H. Thurber and N. Rabinowitz (eds), *Advances in seismic Event Location*, 18, 231-263, Modern approaches in geophysics, Kluwer Academic Publishers, Dordrecht, the Netherlands.

Di Stefano, R., Aldersons, F., Kissling, E., Baccheschi, P., Chiarabba, C., and Giardini D., 2006. Automatic seismic phase picking and consistent observation error assessment: application to the Italian seismicity, *Geophys. J. Int.*, 165, 121-134.

Goldstein, P., Dodge, D., and Firpo, M., 1999. SAC2000:signal processing and analysis tools for seismologists and engineers, *UCRL-JC-135963, Invited contribution to the IASPEI International Handbook of Earthquake and Engineering Seismology.*

Johnson, C.E., Lindh, A.G., and Hirshorn, B., 1994. Robust regional phase association, *U.S.G.S, Open File Report*, 94-621.

Kokinou, E., Papadimitriou, E., Karakostas, V., Kamperis, E., and Vallianatos F., 2006. A review of the Kefalonia Transform Zone (offshore Western Greece) with special emphasis to its prolongation towards the Ionian Abyssal Plain, *Marine Geophysical Researches*, Spinger Verlag, 27, 241-252.

Panagiotakis, C., and Tziritas, G., 2005. A speech/music discriminator based on RMS and zero-crossings, *IEEE Transactions on Multimedia*, 7, 1.

Panagiotakis, C., Grinias, I., and Tziritas, G., 2006. Automatic Human Motion Analysis and Action Recognition in Athletics Videos, *European Signal Processing Conference, Florence*, Italy.

Rowe, C.A., Aster, R.C., Borchers, B., and Young, C.J., 2002. An automatic adaptive algorithm for refining phse picks in large seismic datasets, *Bull. seism. Soc. Am.*, 92(5), 1660-1674.

Wang, J., and Teng, T.-l, 1997. Identification and picking of S phase using an artificial neural network, *Bulletin of the Seismological Society of America*, 87(5), 1140 - 1149.