ELSEVIER

Computers & Geosciences 31 (2005) 1277–1285

COMPUTERS & GEOSCIENCES

www.elsevier.com/locate/cageo

Teseo: A vectoriser of historical seismograms ☆

Stefano Pintore, Matteo Quintiliani*, Diego Franceschi

Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Roma, Italy

Received 15 October 2004; accepted 1 April 2005

Abstract

Historical seismograms contain a rich harvest of information useful for the study of past earthquakes. It is necessary to extract this information by digitising the analogue records if modern analysis is required. *Teseo* has been developed for quick and accurate digitisation of seismogram traces from raster files, introducing a vectorisation step based on piecewise cubic Bézier curves. The vectoriser can handle greyscale images stored in a suitable file format and it offers three concurrent vectorisation methods: manual, automatic by colour selection, and automatic by neural networks. The software that implements the methods described is distributed with open source license.

Keywords: Historical seismogram; Vectorisation; Digitisation

1. Introduction

The use of modern techniques to recover seismological information contained in historical seismograms can supply additional knowledge on past seismicity and ongoing tectonic processes. This important goal is widely acknowledged (Kanamori, 1988; Stein et al., 1988), but it presents some difficulties, since data recorded by early instruments on paper media must be properly processed to obtain numerical data usable for modern analysis. The Sismos Project[1] is aimed at preserving the heritage of historical seismograms owned by all the Italian observatories. For this purpose, an acquisition of the raster images from approximately one million recordings is planned. Data access will be ensured through a dedicated web portal. Digitisation will be performed on a number of the most important events. All this requires a fast and accurate digitisation procedure that is usable on images of large dimensions and independent of the original paper type or recording instrument.

In the digitisation process, seismograms can be classified based on the similarity of the raster images used. Common problems encountered in automatic digitisation of seismic traces are well described by Trifunac et al. (1999). Other recent work on seismogram digitisation includes that by Samardjieva et al. (1998), who created a digital database for historical earthquakes using a manual digitisation process whereby the original records were enlarged by a projector on a screen. Teves-Costa et al. (1999) presented an example of the recovery of source parameters from historical records and developed a semi-automatic method using commercial software on images of 200 dpi resolution. Baskoutas et al. (2000) digitised 1852 seismograms obtained from the Mainka and Wiechert seismographs in the National Observatory of Athens during the period 1911–1960. They developed software usable on black and white images with 1600, or 500 dpi if the image was too large.

☆ Source code is available from server at *http://sismos.ingv.it*

*Corresponding author. Tel.: +39 06 51860671; fax: +39 06 5041303.

E-mail address: quintiliani@ingv.it (M. Quintiliani).

[1] Istituto Nazionale di Geofisica e Vulcanologia.

0098-3004/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.cageo.2005.04.001

The digitisation procedure usually involves:

(i) extraction of the sample sequence directly from the image, in a manual or automatic way,
(ii) correct mapping from the (x, y) image coordinates to the amplitude and time of the samples.

We present here a different method that relies on an intermediate parametric vectorial representation of the seismogram trace using piecewise cubic Bézier curves. To implement this method, a software tool named *Teseo* has been developed. It offers one manual and two automatic modes of operation. The first automatic mode is analytical and the second is based on neural networks.

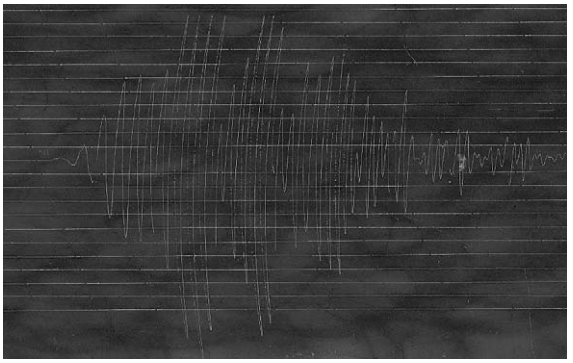

Fig. 1. Example of seismogram curvature derived from use of a tracing needle mounted on a finite-length pivoting arm.

2. Data

Analogue seismograms recorded on paper result from the response of a seismometer and a recording system to ground motion. The main classes of traditional seismometers include short-, intermediate- and long-period instruments. The recording system for analogue instruments is characterised by several mechanisms, the most important being the kind of support used (smoked, photographic, thermal), the type of tracing device (needle, light beam) and the paper speed. The digitisation of traces contained in such seismograms is the final objective. The Sismos Project has set scanning specifications to ensure that all information contained in the paper seismograms are preserved, while considering both the characteristics of the support and the expected digital output. To this end, it has been found that the raster images must be acquired using very high-quality A0 scanners, at a resolution of 1016 dpi with 256 grey levels for the raster images containing the earthquake traces. The standard format used to store these images is plain TIFF[2] (Adobe Developers Association, 1992). This choice requires approximately 400–500 MB for a sheet of paper measuring 120 cm × 40 cm and it guarantees integrity and consistency of the information contained in the raster image. *Teseo* software is designed to offer a tool for vectorisation of all classes of seismograms acquired by the Sismos project. Obvious difficulties, however, follow from the different types of seismogram objects for the vectorisation. Most problems are related to the quality of the trace recorded on the paper.

It is possible to delimit various main blocks of problems or cases.

Fig. 1 shows a seismogram produced by a Wiechert seismograph on smoked paper. The curvature of the trace resulting from the needle mechanism is evident. In this case, there is a loss of correspondence between the abscissa and time, because the trace at its maximum amplitude is somewhat ahead of the zero crossing at the same time. This justifies the use of a parametric representation of the seismogram trace.

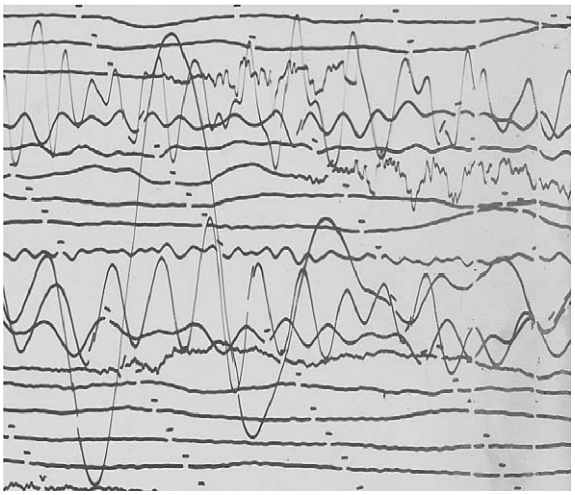

Fig. 2. Example of a seismogram trace crossing on a photographic record. Trace crossing occurs in any type of drum recording medium.

Crossing traces as shown in Fig. 2 are another problem when digitising. In fact, digitisation systems based only on the determination of trace colour are not able to distinguish between points belonging to one trace rather than the another. Sometimes this task can be very difficult, even for a skilled seismologist.

When digitising, it is also important to consider the thickness of the traces in relation to the frequency of the signal. An algorithm calculating the weighted mean of

[2]Tagged Image File Format.