Teseo²: Turn the Eldest Seismograms into the Electronic Original Ones

I.N.G.V.

Stefano Pintore and Matteo Quintiliani

Istituto Nazionale di Geofisica e Vulcanologia Roma, Italy

Introduction

Historical seismograms contain a rich harvest of information useful for the study of past earthquakes. This requires proper digitization of the analog records if modern analysis is sought. The digitization procedure usually involves the extraction of the sample sequence directly from the image and, successively, a correction mapping from the (x,y) image coordinates to the amplitude and time of the samples. We present here a different digitization approach that relies on an intermediate parametric vectorial representation of the seismogram trace using piecewise cubic Bézier curves. Our proposed workflow standardizes the historical seismograms digitization process into various stages insuring proper quality control.

We have developed a software for seismogram digitization/vectorization named Teseo² [Pintore et al., 2005] in the framework of the Sismos project [Michelini and Sismos Team, 2005]. Teseo² is a plug–in for GIMP – a multiplatform photo manipulation tool [Gimp, 2005] – that extends its functionalities for seismological studies.

Historical Seismograms

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. Figures 1–2 show two different kinds of historical seismograms. The digitization of traces contained in such seismograms is the final objective.

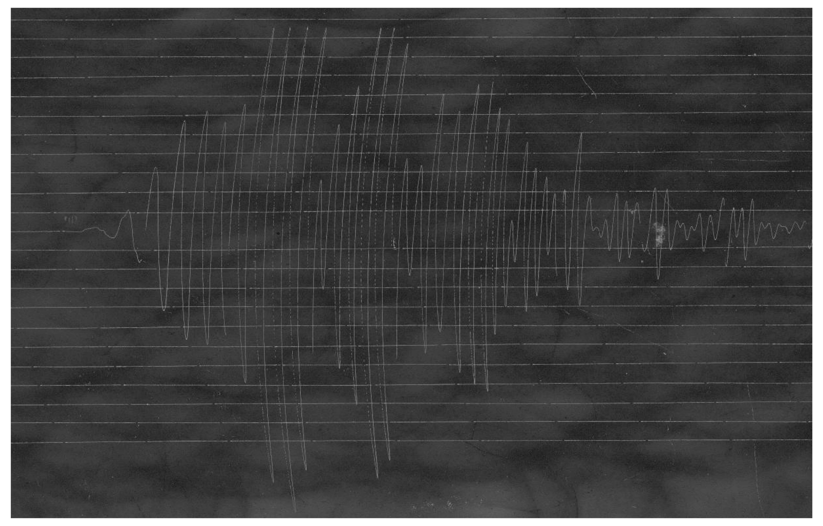


Fig. 1. Example of a seismogram produced by a Wiechert seismograph on smoked paper. The curvature of the trace resulting from the needle mechanism is evident. In this case, there is a loss of correspondence between the abscissa and time, because the trace at its maximum amplitude is somewhat ahead of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

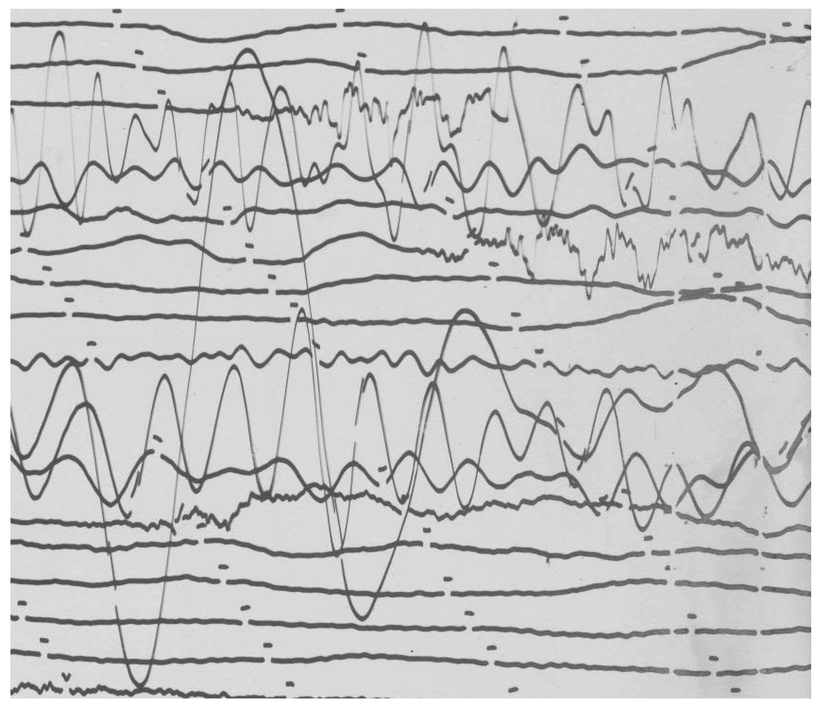
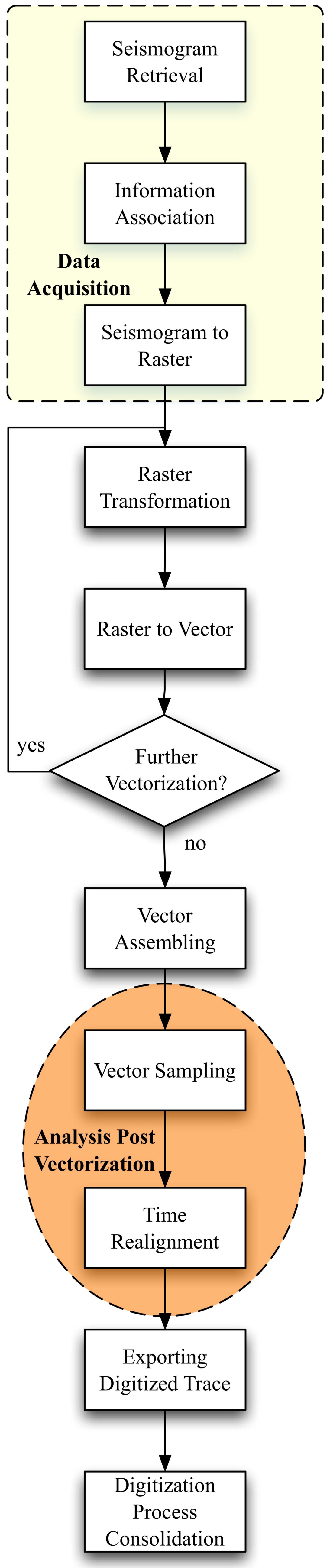


Fig. 2. Example of a seismogram trace crossing on a photographic record. Trace crossing occurs in any type of drum recording medium.

Seismogram Digitization Workflow



We propose a complete digitization workflow that permits the quality control of the data produced.

- *Seismogram Retrieval*: obtaining the seismogram paper.
- *Information Association*: finding out about station location, instrument parameters, earthquake bulletins, etc. This kind of information is sometime indispensable in the Analysis after the trace vectorization.
- *Seismogram to Raster*: in order to preserve the information contained in the paper seismograms, it is better to acquire it by a high–quality scanner. The choice about resolution and color–depth mostly depends on the instrument that recorded the trace, the sampling of the final signal expected, and the seismogram characteristic such as trace thickness.
- *Raster Transformation*: enhancing the "readability" of a seismogram is very recommended before vectorizing it. The most useful filters are those help you to increase the brightness and contrast of the image.
- *Raster to Vector*: this intermediate step in digitization process allows you to reproduce the shape of the trace using a vectorial representation. Usually there are discontinuities in the signal trace recorded on the image that force the user to create iteratively multiple vector pieces.
- *Vector Assembling*: ordering vector pieces and filling the gaps.
- *Vector Sampling*: extraction of the points sequence belonging to the vector.
- *Time Realignment*: assignment of the right time and amplitude to each point. Curvature correction could be required.
- *Exporting Digitized Trace*: saving the sequence of samples choosing an appropriate file format.
- *Digitization Process Consolidation*: saving all the information related to the digitization process in a single place. This permits quality control.

Implementation

The data acquisition is very expensive in terms of time spent to find out seismograms and associated information. Moreover, high–quality and big dimension scanners require considerable investments in hardware and technical staff. This can be achieved in the framework of large scale projects such as Sismos.

Images produced by Sismos have usually a resolution of 1016 dpi with 256 grey levels. The standard format used to store these images is plain TIFF. This choice requires approximately 300–400 MB for a sheet of paper measuring 120 cm $\times$ 40 cm.

In order to implement the rest of the digitization workflow we developed Teseo², a plug–in for GIMP. GIMP offers many features and capabilities for manipulation of images regardless of their size and it is an expandable and extensible open–source software.

Teseo² relies on GIMP only for the raster transformation, the vectorial representation of the seismic trace and the consolidation of the digitization process. This last step is achieved putting all information related to the digitization process into an *xcf* file. The *xcf* format – the GIMP proprietary format – can save different raster layers and vectors produced as well as arbitrary pieces of data. Moreover, *xcf* natively supports gzip and bzip2 compression.

Teseo² is also developed following the "Open–Source" philosophy and it is freely distributed under GPL license. Finally, it is cross–platform and the sources, the binaries for Linux, Windows and Mac OS X, are periodically updated on the Sismos web site.

Teseo²

Teseo² allows primarily for:

- additional operations on the vectorized trace (i.e. resampling and alignment)
- supervised vectorization algorithms (color weighted mean)
- analysis after trace vectorization, such as curvature correction and time realignment
- trace import/export in several formats (such as SAC, SVG, DXF, ASCII, Timemarks distances).

In order to keep track of the stages and parameters of a seismogram vectorization, Teseo² is able to write this information into the image saved in *xcf* format.

Raster Transformation

GIMP offers a variety of filters and instruments to manipulate images. At the moment, Teseo² provides a graphical filter useful to "clean" a seismogram. What do we intend to clean a seismogram? Before vectorizing it is advantageous to remove horizontal traces crossing it while mantaining trace continuity. The main idea is to fill unwanted horizontal or vertical segments with the background color of the seismogram. Figure 3 shows an example.

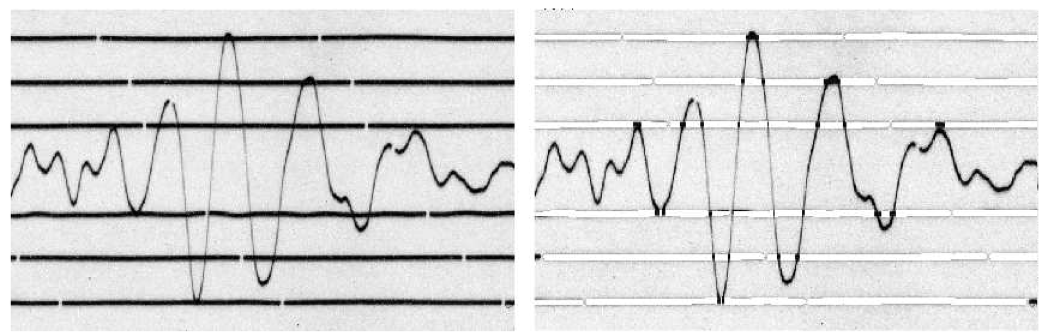


Fig. 3. A particular of a trace before and after applying the "Clean" filter.

Raster to Vector

You can manually vectorize the traces by GIMP Path tool creating several piecewise cubic Bézier curves or polylines. In figure 4 is shown an example of a cubic Bézier curve.

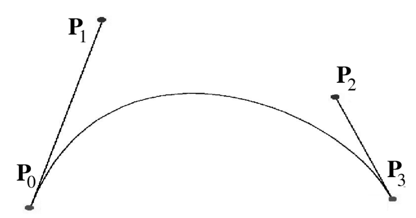


Fig. 4. A graphical representation of a cubic Bézier curve. $\mathbf{P}_0$ and $\mathbf{P}_3$ are named anchor points, and $\mathbf{P}_1$ and $\mathbf{P}_2$ are named control points.

Teseo² is designed to easily add algorithms for automatic seismogram vectorization. An iterative procedure takes place whereby at each step the algorithm is executed providing it with:

- a rectangular portion of the image centred at the last point of the current path;
- information regarding the closest previous points;
- a preferred direction suggested by the user clicking on the arrow buttons showned in figure 5;

in order to find the next point.

Presently, Teseo² uses an algorithm based on a weighted mean of the trace color that is widely described in [Pintore et al., 2005].

In future versions of Teseo² more algorithms should be available (neural network approach too).

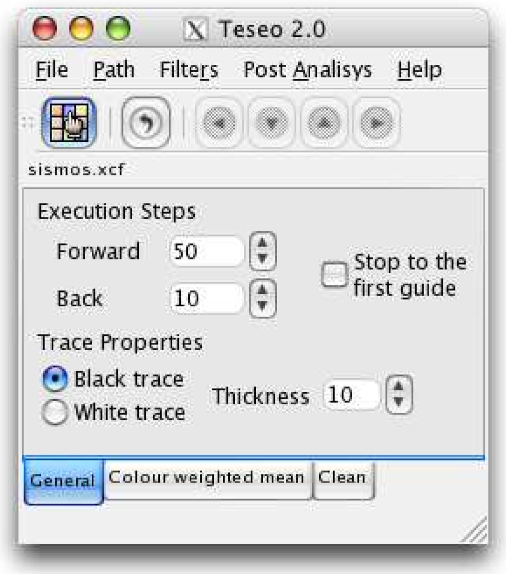


Fig. 5. Snapshot of the Teseo² main window.

Time realignment

The seismogram curve on the image has to be corrected to become a seismic data with right amplitude and time. There are many errors that could be introduced during the digitization procedure that msut be taken into account. The algorithm we use was originally written in Fortran by A. Schlupp. The algorithm needs some parameters, for a few of them Teseo² offers some other instruments to evaluate. The algorithm is described in a personal comunication provided by A. Schlupp which is going to be published. It is based on the original Cadeck formula [Cadeck, 1987].

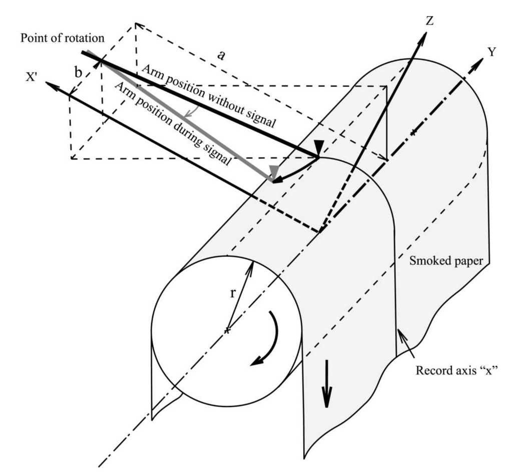


Fig. 6. Recording system model used by the curvature correction algorithm.

Exporting Digitized Trace

Although Bézier curves allows an unlimited level of detail in resampling we limit the sample rate to one sample per pixel.

Contact

Official web site: `http://sismos.ingv.it/teseo/`
Developer e–mail: `teseo@ingv.it`
User mailing–list: `teseo-user@yahoogroups.com`

References

[Cadeck, 1987]Cadeck, O. (1987). Studying earthquake ground motion in prague from wiechert seismograph records. *Gerl. Beitr. Geoph.*, 96:438–447.

[Gimp, 2005]Gimp (2005). *GNU Image Manipulation Program User Manual*. The GIMP Documentation Team. URL `http://docs.gimp.org/`.

[Michelini and Sismos Team, 2005]Michelini, A. and Sismos Team (2005). Collection, digitization and distribution of historical seismological data at ingv. *EOS*, 86(28).

[Pintore et al., 2005]Pintore, S., Quintiliani, M., and Franceschi, D. (2005). Teseo: a vectoriser of historical seismograms. *Computers & Geosciences*, 31(10):1277–1285.

Contact information: Matteo Quintiliani, Via di Vigna Murata, 605 – 00143 – Roma – Italy; Phone: +39 06–51860–671; Email: *quintiliani@ingv.it*; Web: *http://sismos.ingv.it/teseo/*