AGU 2021, Fall Meeting, Supplement to Poster S25E-0298:

Fidelity of Legacy Seismic Data on Microformats

Thomas Lee[1], Miaki Ishii[1], and Paul Okubo[2]

[1]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA
[2]School of Ocean and Earth Science Technology, University of Hawai‘i at Mānoa, Honolulu, HI

Background

Recording on paper was common throughout the first hundred years of modern seismology. Many of these records are still extant today, and moreover, many exist in both original paper form and as microform copies (i.e., have been replicated from original paper onto microfilm reels or microfiche chips). When these paper records were filmed onto microformats, an extra step was introduced into the processing of the data that may have effects on fidelity (Figure 1). Therefore, it is crucial to understand what differences may exist between digitizations derived from microform and from paper records in order to prioritize data rescue in an informed way.

Data Details

The records displayed on the poster and in this supplement are from July 15, 1945 and show the world's first nuclear test, Trinity (Figure 2), as recorded by the short-period vertical instrument in Tucson, Arizona. The paper record was scanned using a high-resolution flatbed color scanner. The microform record was scanned from 15mm microfilm, and was originally filmed as a part of the USGS Historical Seismogram Filming project. Due to differences in available scanning technologies for these media, the paper scan has a resolution of about 1000dpi and is done in color, whereas the microfilm scan has a resolution of about[1] 500dpi (relative to the original) and is in grayscale. These coarser pixels can affect the amplitude of traces overall. Additionally, the microfilm scan must be done in two parts, yielding a left side and right side scan that must be aligned at some point in the digitization process[2].

Investigation of Tonal Depth

The data which is recoverable from a given scanned record during digitization depends upon how visible waveforms are. Of particular concern are areas where high-amplitude waveforms are recorded. This is because while these are generally portions of the record data-users are interested in – such as waveforms associated wit nuclear explosions and earthquakes –, the trace is likely faded out due to rapid movement of the armature. Indeed, upon initial inspection, these parts of the paper scan show significant thinning of the trace, however waveforms are still visible (***A*** on poster). On the other hand, the microfilm often

[1]Because the microfilm is a 15mm reel and is enlarged using a series of lenses before scanning, the typical direct relationship of digital-units-per-length-units is not applicable. Therefore, we estimate the dpi by comparing the number of pixels in the paper scan with those in the microfilm scan and assuming the size of the original record.

[2]The poster displays the right side microfilm scan (*red false color*) overlaid over the corresponding part of the paper record (blue false color).

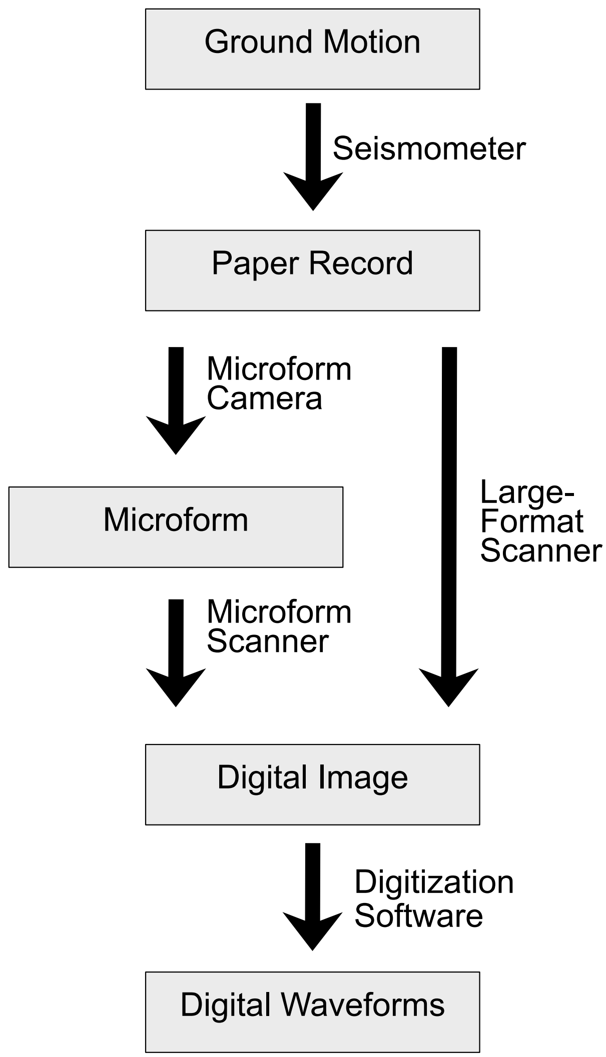

Figure 1: Flowchart showing the path of information from initial ground motion to digital waveforms. Gray boxes show type of information, and black arrows indicate tools for processing this information.

Figure 2: Trinity nuclear test, July 15, 1945 (*Source: lanl.gov*)

seems to have only the peaks and troughs of the waveforms with the information between these points missing altogether.

These different behaviours can be confirmed and better understood with analysis of the grayscale tones present in both images. Tonal depth can be understood as the number of distinct tones which are present in an image. Both images are in 8-bit grayscale (ranging from black at 0 to white at 255), however, not all of these tones are utilized. In the paper record, pixels cover a large range of tones (roughly from $\sim$70 to $\sim$220), on the other hand, the microfilm record is essentially bi-tonal with pixels only around 0 and 255 (Figure 3).

A transformation can be solved for and applied to these images that re-assigns grayscale values of pixels to try and match a given distribution while preserving relative relations (if pixel A is darker than pixel B, then pixel B will not be darker than pixel A after transformation). The paper record can be transformed in this way to match the microfilm, but an image with the tonal distribution of the microfilm cannot be transformed to replicate the paper image (Figure 4). This makes sense when compared to the tone-distribution-curves that show an essentially bi-tonal microfilm distribution and a more varied paper distribution. The transformation of the paper record to replicate the microfilm essentially bins the data into two boxes, at which point there is no-longer the depth of information present to recover the variance which the paper image has.

The results of this analysis show that there are no middle-ranged "gray" pixels in the microfilm scans, there are only black and white pixels. The implications for this with regard to digitization are that the microfilm will not be able to display the information between the peaks at large armature swings. The thinning of the lines themselves due to the lower exposure time (in the case of galvanometer photo-records) is what makes this crucial information essentially invisible on microfilm scans.

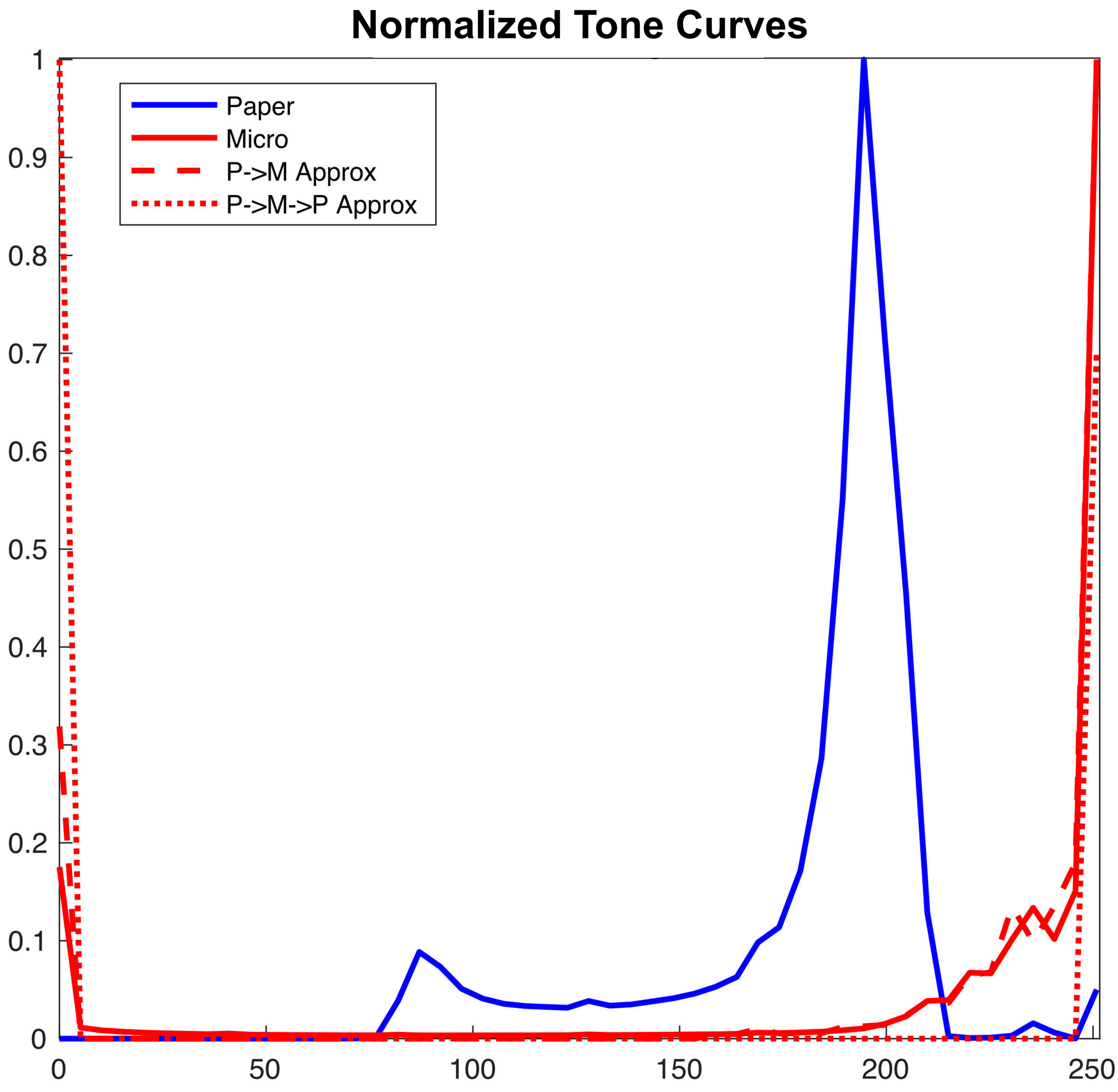

Figure 3: Normalized distributions of tones from 0 to 255 for the paper record (*solid blue curve*), the microfilm record (*solid red curve*), the paper record transformed to replicate the microfilm (dashed red curve), and that record transformed again to recover the paper tonal distribution (*dotted red curve*).

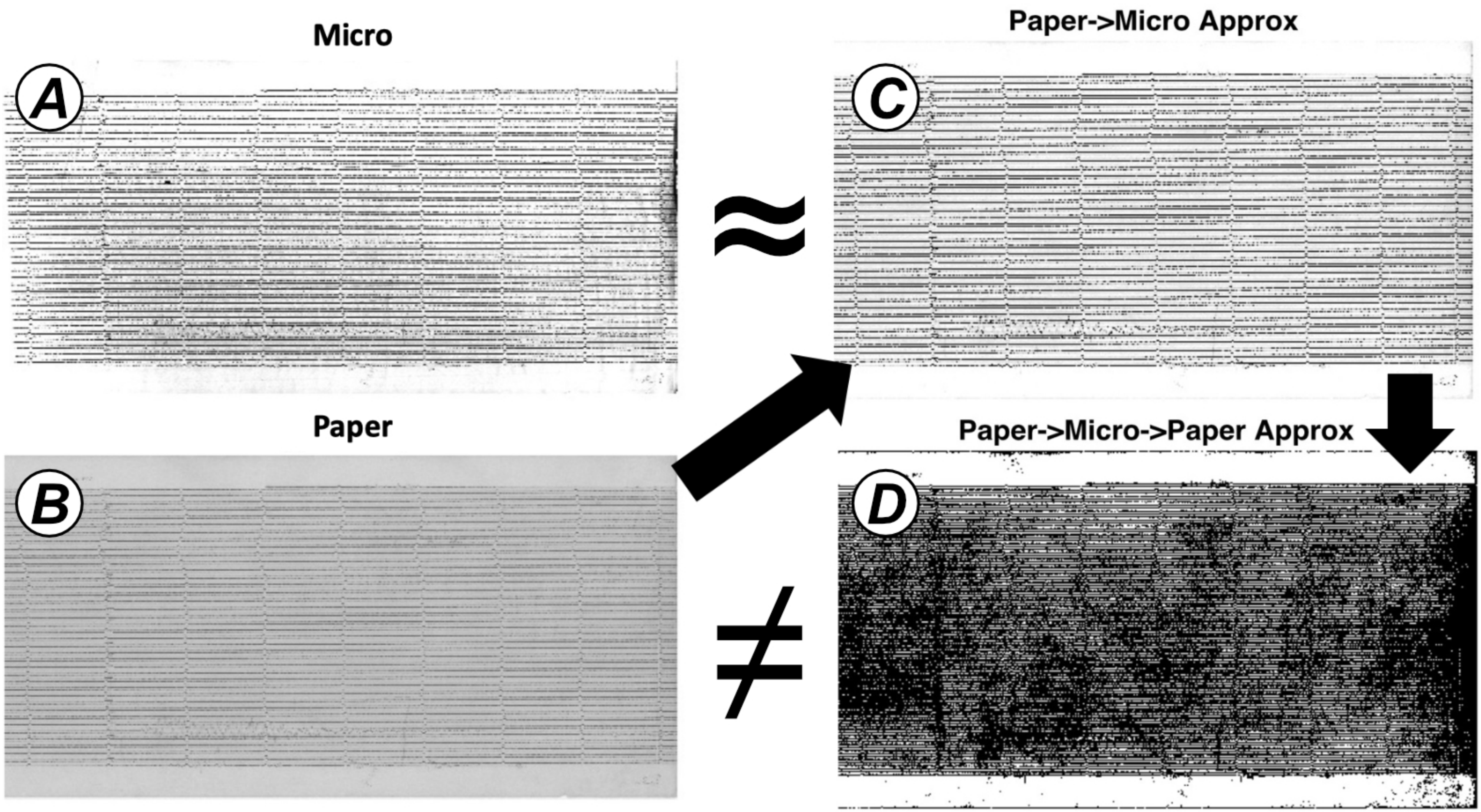

Figure 4: (*A*) Scanned image from left side of microform. (*B*) Scanned image from paper image showing the same part of seismogram as the microform. (*C*) Paper image with filter applied to match the tonal distribution of the microform image. (*D*) Microform approximation from paper filtered once again to try match the tonal distribution of the original paper image.

Links of Interest

- http://www.seismology.harvard.edu/research/DigitSeis.html
  Information on DigitSeis, the seismogram-digitization software used in this project.
- http://www.seismology.harvard.edu/HRV/archive.html
  Archives of historical seismograms from the Harvard, Massachusetts station (HRV).
- http://www.seismology.harvard.edy/research/microform/TrinityTimeseries.jpg
  Image showing the timed-waveforms from the Trinity test as recorded in Tuscon[3]. Note that waveforms are now aligned at minute marks, but there is significant non-linear distortion between minute marks.

[3]The waveforms in ***B*** on the poster are showing x and y position that are obtained after digitization but before timing