Seismology with Dark Data: Image-Based Processing of Analog Records Using Machine Learning for the Rangely Earthquake Control Experiment

by Kaiwen Wang, William L. Ellsworth, Gregory C. Beroza, Gordon Williams, Miao Zhang, Dustin Schroeder, and Justin Rubinstein

ABSTRACT

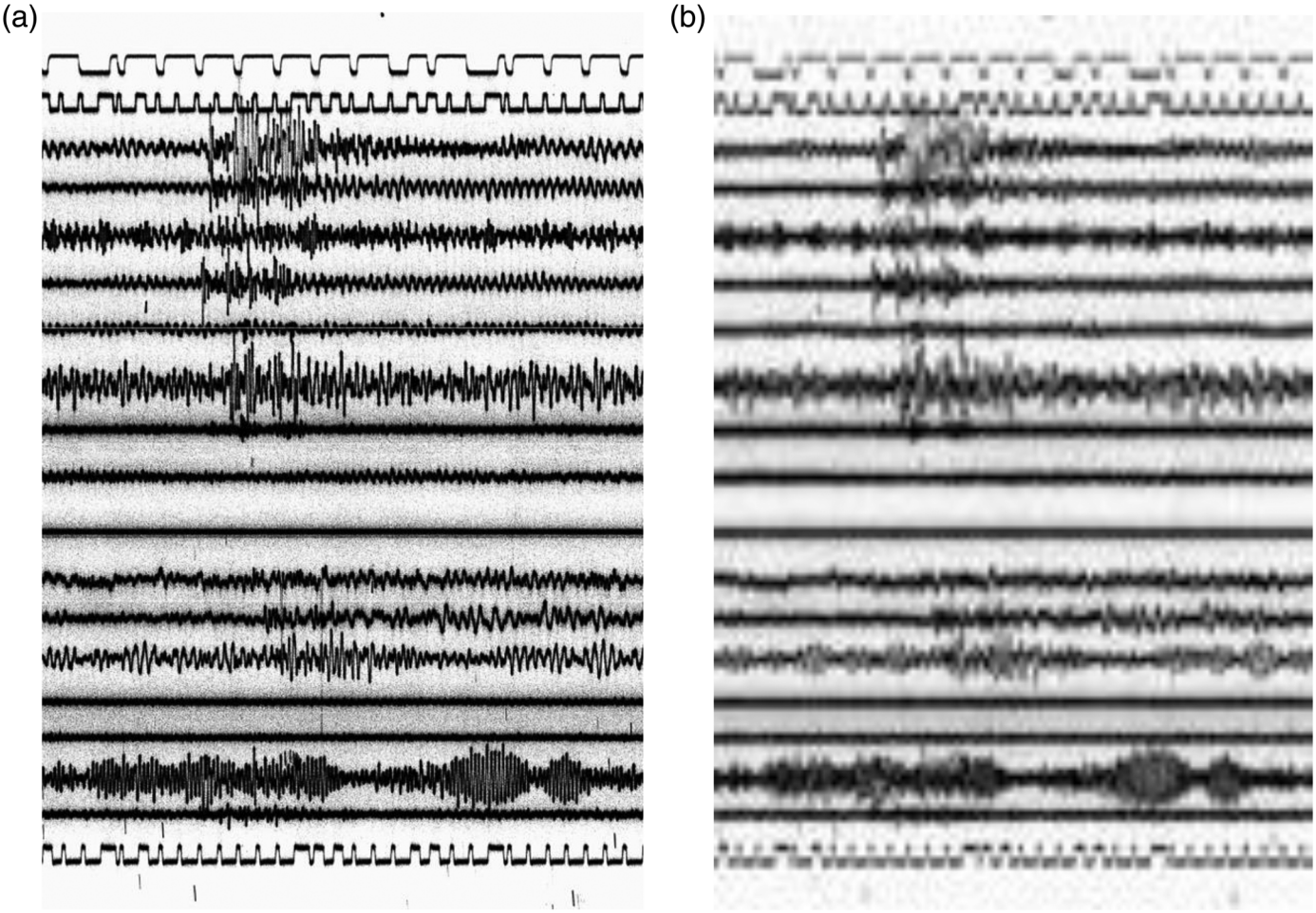
Before the digital era, seismograms were recorded in analog form and read manually by analysts. The digital era represents only about 25% of the total time span of instrumental seismology. Analog data provide important constraints on earthquake processes over the long term, and in some cases are the only data available. The media on which analog data are recorded degrades with time and there is an urgent need for cost-effective approaches to preserve the information they contain. In this study, we work directly with images by constructing a set of image-based methods for earthquake processing, rather than pursue the usual approach of converting analog data to vector time series. We demonstrate this approach on one month of continuous Develocorder films from the Rangely earthquake control experiment run by the U.S. Geological Survey (USGS). We scan the films into images and compress these into low-dimensional feature vectors as input to a classifier that separates earthquakes from noise in a defined feature space. We feed the detected event images into a short-term average/long-term average (STA/LTA) picker, a grid-search associator, and a 2D image correlator to measure both absolute arrival times and relative arrival-time differences between events. We use these measurements to locate the earthquakes using hypoDD. In the month that we studied, we identified 40 events clustered near the injection wells. In the original study, Raleigh *et al.* (1976) identified only 32 events during the same period. Scanning without vectorizing analog seismograms represents an attractive approach to archiving these perishable data. We demonstrated that it is possible to carry out precision seismology directly on such images. Our approach has the potential for wide application to analog seismograms.

INTRODUCTION

In September 1969, the U.S. Geological Survey (USGS) began a controlled injection experiment in an oil field in Rangely, Colorado, to test the effective stress hypothesis: increased fluid pressures lower the effective normal stress, lowering the friction on the fault and in turn increasing the likelihood of earthquakes (Raleigh *et al.*, 1976). This experiment is unique in that it is the only human-controlled injection experiment intended to trigger earthquakes. Before, during, and after the period when fluid pressure was manipulated, a telemetered network of 14 stations recorded the earthquake activity. Unfortunately, almost the entire earthquake catalog (except a couple days reported by O'Neill and Healy, 1973 and Gibbs *et al.*, 1972) for this experiment has been lost; however, the original 16 mm Develocorder films are still available.

Before the digital era in seismology, analog seismographs were widely used and recorded decades of seismic data stretching back to the late nineteenth century. Many historical seismograms are properly stored and are still in remarkably good condition, but archiving, storing, and analyzing these data present a continuing challenge. The USGS, along with other institutions, preserved tens of thousands of analog films that contain a unique record of continuous ground motion (Ishii *et al.*, 2014). Recovering information from these precious analog seismograms should be a priority. In recent years, several ambitious attempts have been made to digitize analog seismograms from images (Baskoutas *et al.*, 2000; Pintore *et al.*, 2005; Benson *et al.*, 2011; Xu and Xu, 2014; Bogiatzis and Ishii, 2016; Wang *et al.*, 2016). These efforts focused on following and vectorizing traces to create time series. When traces overlap as often happens during earthquakes, manual intervention is usually required to follow signals of interest. This can be difficult or impossible. Machine vision and processing techniques offer a different path. In this study, we skip the time series digitization step and propose a set of image-oriented detection and location methods to process images of the analog seismograms. Our goal is to build a set of image-based approaches to extract information from the microfilm data, so that we can revisit this unique and valuable experiment.

doi: 10.1785/0220180298



▲ **Figure 1.** An example of scanned microfilm. (a) High-resolution image used for precise timing. (b) Downsampled image used for event detection and initial processing.

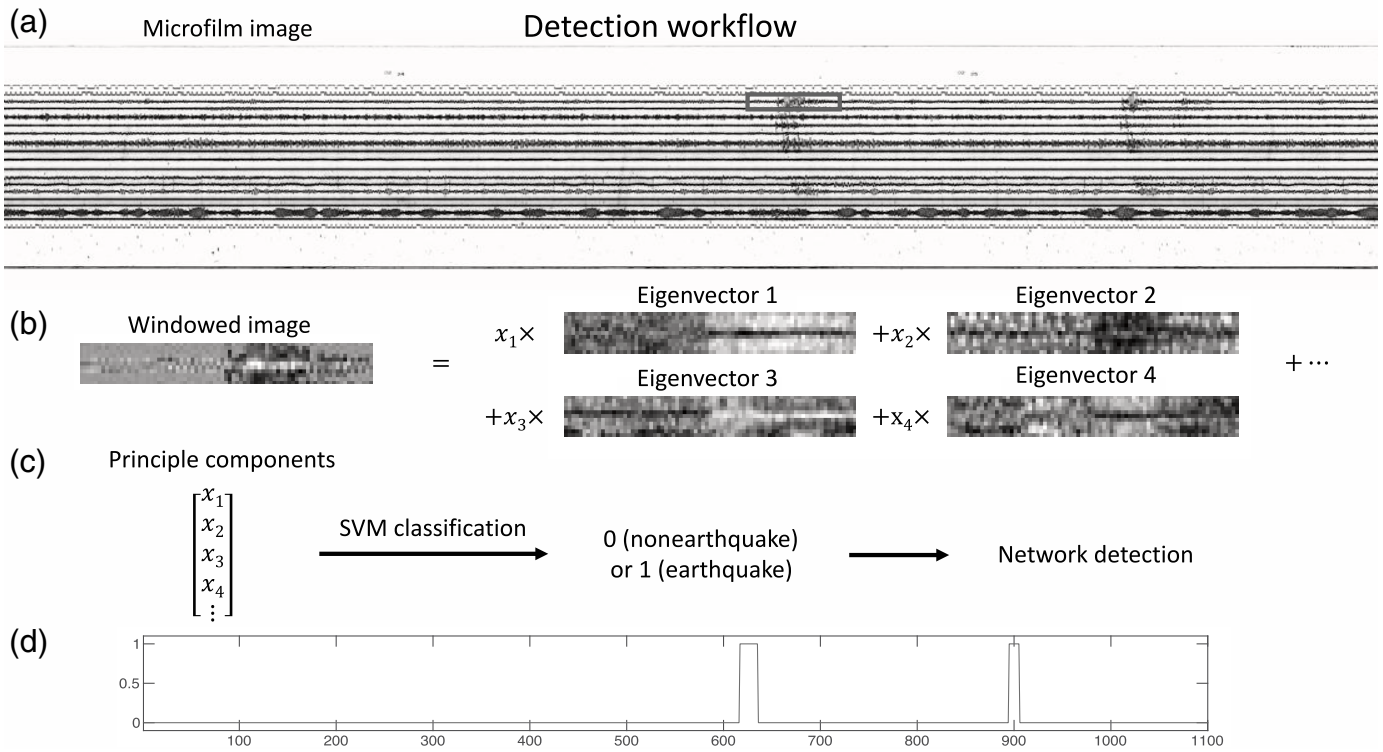
METHODOLOGY

In this section, we describe our automatic processing workflow to construct a local earthquake catalog from Develocorder films. It contains two major components: earthquake detection and earthquake location. Each day-long Develocorder film is scanned into ~700 images that total about 10 GB in size (Fig. 1). To detect events in this large and unlabeled image dataset, we first apply an unsupervised learning algorithm—principle component analysis (PCA, Jolliffe, 2011)—to represent image features in lower dimensions. We then use a support vector machine (SVM, Cortes and Vapnik, 1995) classifier to represent earthquake events in the feature space defined by PCA. To locate these detected events, we first pick their absolute arrival time by a short-term average/long-term average (STA/LTA) picker that runs directly on the scaled image and its horizontal gradient. The picks that result are fed into a grid-search associator to form individual earthquakes that are input to a location program. We apply 2D image correlation around the arrival time to measure differential arrival times between similar events. These are used as input for double-difference relocation. We carry out double-difference location using both the absolute arrival time and the cross-correlation time differences to determine precise earthquake locations.

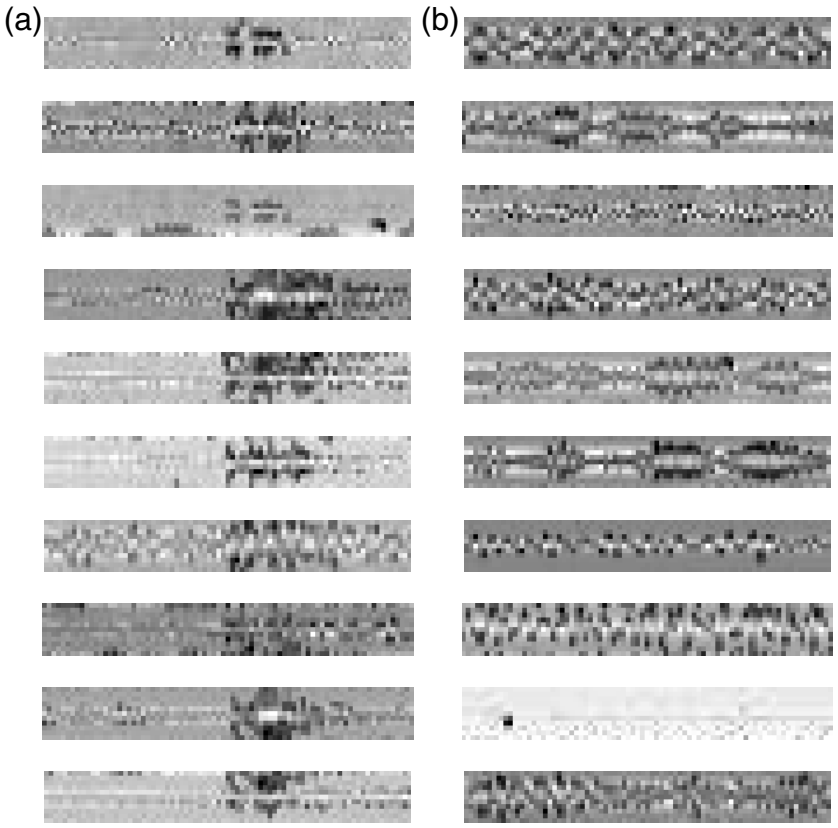
Detection

Figure 2 shows an overview of the detection workflow. We first segment each image into overlapping windows centered on traces and remove the mean image. We then build a training set with half-earthquake and half-nonearthquake examples for PCA and SVM binary classification. The detections on each channel are combined to give network detections. The final outputs of the detection workflow are images that contain the signals we are interested in—local earthquakes in the case of the Rangely dataset.

We scan the microfilms into RGB color images of 140 s length with 10 s overlap on each end. A sample image is shown in Figure 1. The top two traces and the bottom trace are time codes. The middle 16 traces correspond to recordings from 14 seismic stations: 14 vertical components and 2 low-gain channels. The images are originally $5880 \times 22,000$ pixels, which are 15 MB jpeg files. For computational efficiency, we downsample the images to 252×1100, which provides sufficient resolution for event detection. The detection step is conducted on the



▲ **Figure 2.** Detection workflow: (a) Image segmentation and preprocessing; (b) feature extraction by principle component analysis (PCA); (c) support vector machine (SVM) binary classification; and (d) combine network detections.

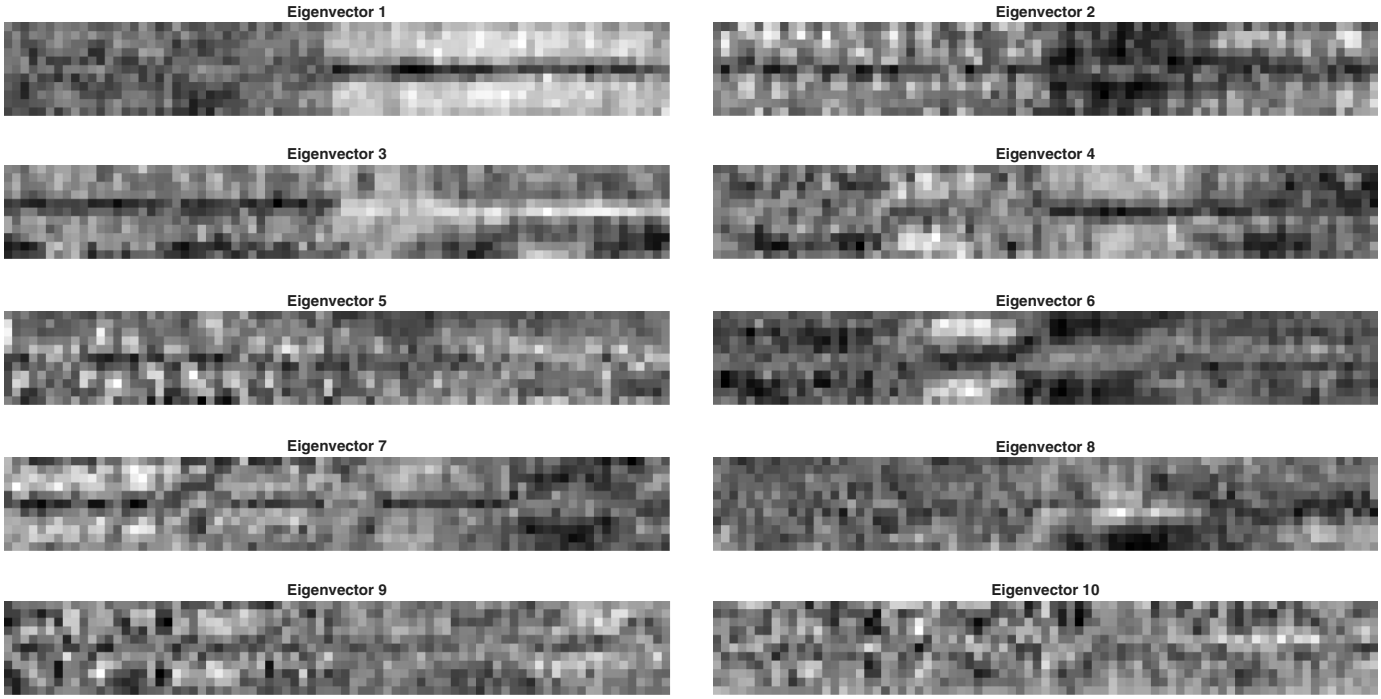


▲ **Figure 3.** (a) Earthquake and (b) nonearthquake examples in the training set.

downsampled images to identify candidate local earthquakes. In the location processing that follows, we switch back to the high-resolution images to ensure location accuracy.

We preprocess and divide the downsampled images into segments of single-channel waveform images before running the classifier on moving windows. We first convert the RGB image into gray scale and stack horizontally to find the trace center positions. We then take 10 s overlapping windows centered on the 16 channels with fixed height and 0.1 s lag. The mean image of each trace is removed to eliminate the differences in light conditions of the films and background noise levels.

Because we have a large and unlabeled dataset, we first apply an unsupervised learning PCA algorithm to extract image features and reduce data dimensions. We follow the eigenface algorithm for facial recognition (Turk and Pentland, 1991) to construct an eigen earthquake representation of earthquake images. To learn the feature space, we build a small dataset of 41 earthquake examples and 40 nonearthquake examples by manually inspecting the films. As can be seen in Figure 3, earthquake examples share similar patterns so that PCA finds a set of eigenvectors to represent the high-dimension



▲ **Figure 4.** Top 10 eigenvectors that define the earthquake feature space.

image space in a lower dimensional subspace. To prevent overfitting to the small training set used for classification, we choose the top 10 eigenvectors to form the feature space (Fig. 4). The learned feature space is used to project new windows of images to obtain low-dimensional vector representations of them. With more earthquake examples, the feature space can be updated or recalculated on a larger training set, and the performance of the classifier should improve.

We next train an SVM classifier that takes the 10-element vector from the 10 eigen components as input and outputs the binary classification of whether or not the corresponding window contains earthquakes. We split our manually labeled dataset into a training set (50%) and a test set (50%). An SVM classifier is trained on the training set and its performance is evaluated on the test set. On the test set, we have a false negative rate of 0.17, a false positive rate of 0.18, a true positive rate of 0.82, and a true negative rate of 0.83. The accuracy is above 80%, showing that our baseline model produces an acceptable result even with a rather small training set. As mentioned previously, the method should improve with a larger training data set.

To lower the false detection rate further, we associate single-channel detection results to check detection across the network. The predicted labels among the 16 channels are combined and events with less than three activations within 10 s are rejected. Figure 5 shows one of the network detection outputs overlain on the original microfilm image. The detection results clearly show two events in the 2-min-long image. Although the classifier misses some activation, for example, the second event in the fourth channel, it successfully classifies most of the activated windows with a manageable false positive rate. We also note that there is a small event right after the second event, which increases the detection score on some channels; however, it will be difficult to detect such a weak event reliably with the downsampled image resolution. We expect that further refinement of the baseline model would allow detection of weaker events.

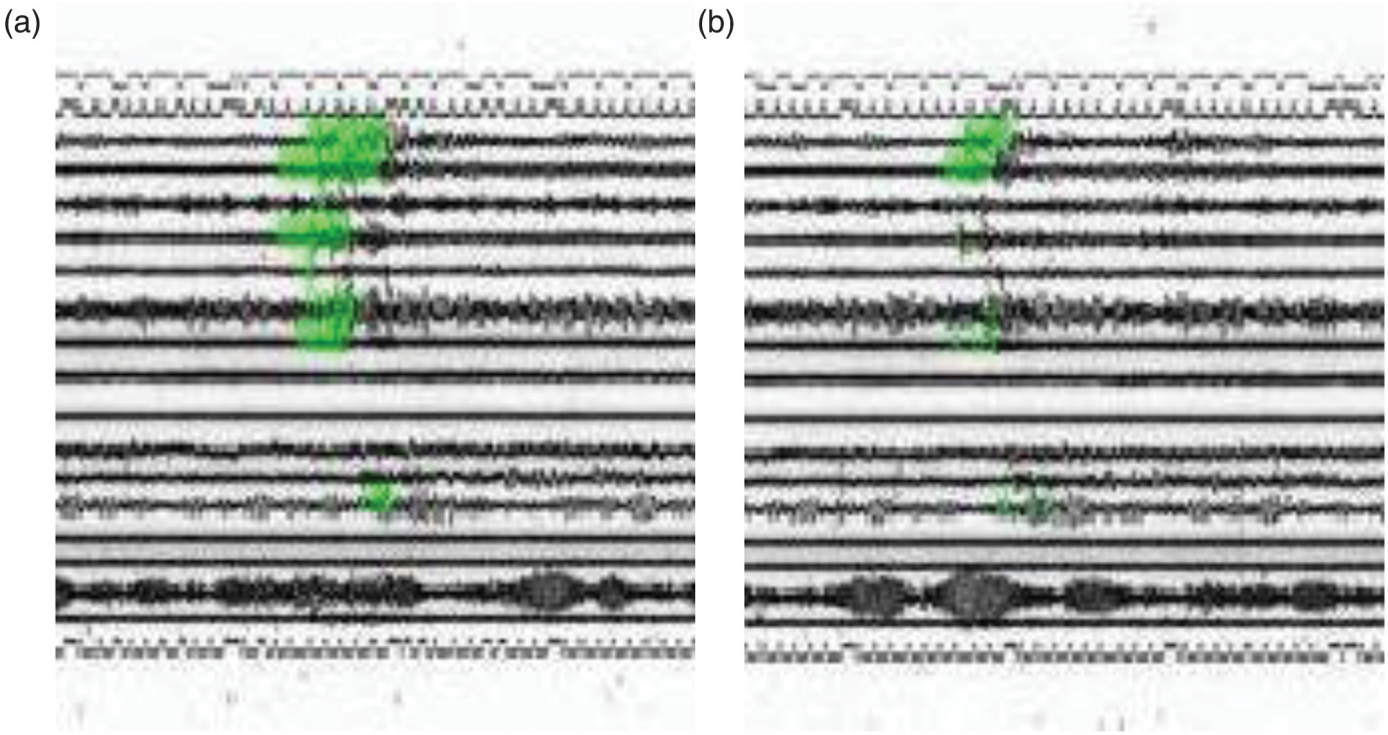
Location

The next step in processing is to use the images selected as events in the initial step for earthquake location. We use both the absolute travel time of the P phases and relative time differences between events in the location step. The two kinds of travel-time measurements are input to the double-difference location program hypoDD (Waldhauser and Ellsworth, 2000; Waldhauser, 2001). To increase timing accuracy, we make skew corrections to the films and match the channels on the films to the stations.

We adapt the STA/LTA picker (Allen, 1978) to run on the 2D image. Before picking, we check the symmetry of the trace around its baseline and mask the strong asymmetric signals to reduce the overlapping energy from adjacent traces and electronic noise spikes. This preprocessing step efficiently lowers the false positive rate of the picks. The characteristic function $E(t)$ (Fig. 6) for the STA/LTA picker is defined as:

$$E(t) = Cf(t)^2 + f'(t)^2, \quad (1)$$

in which $f(t)$ is the summed pixel density of a single-channel image slice in each time increment, $f'(t)$ is the summed pixel density of the horizontal gradient of the single-channel image in each time increment, and C is a weighting constant that controls the relative contribution of the two terms. The first term will help pick events that have clear first arrivals. A complication of the recording system is that when the amplitude



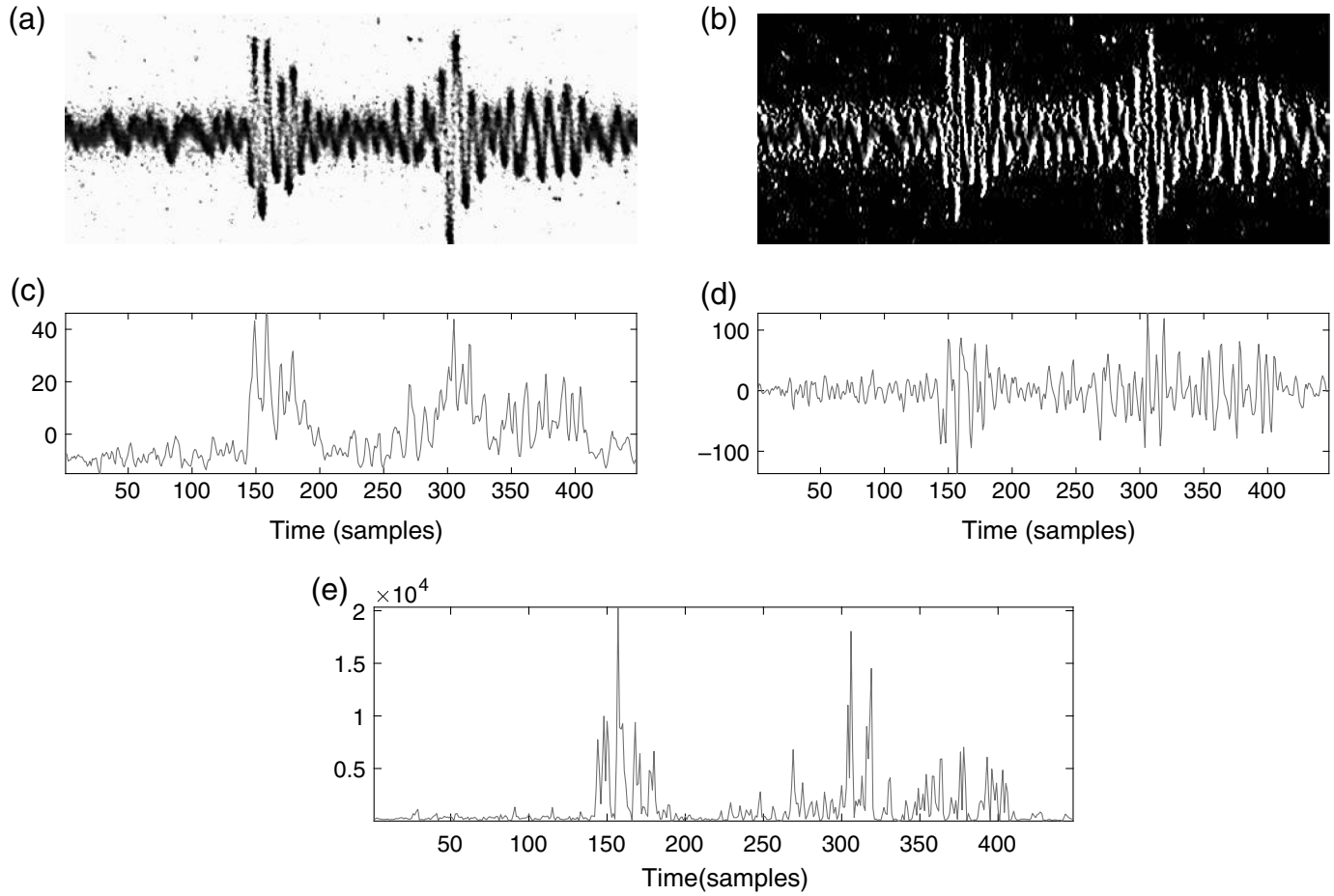
▲ **Figure 5.** Detections overlaid on original image. The green pixels show the center of the windows with positive detections. (a) Zoom in view on the first event and (b) zoom in view on the second event and the following missed event.

gets larger during an earthquake, the light source is moving more quickly and the film is more lightly exposed, making the trace more difficult to follow. In such conditions, the horizontal gradient has a sharp increase at the time when the trace suddenly changes. Varying C adjusts the sensitivity to these two situations. Here, we choose $C = 2$ after checking the picking results. The single-channel picking results are fed into a grid-search associator that calculates possible moveout patterns to test whether the picks are consistent with candidate locations. The associator rejects the bad picks, electronic noise, and events that are not local to the Rangely oil field.

To increase location accuracy, we compute 2D cross correlations between the images of the two similar events. We select a window with 1 s before and 2 s after the *P*-wave arrival based on the picked time. For the channels without a pick, we calculate the theoretical arrival time and choose the window accordingly. Figure 7 shows the processing steps for image correlation. We preprocess the windowed images by muting the overlaps as described previously. We also estimate the noise levels of each channel (Fig. 7b) and remove the center slice of the image according to the amplitude of the noise (Fig. 7c) so that the correlation discounts the contribution from the noise. Following that, we compute the 2D cross correlation between windows from different events on the same channel. We scale the height of the windows to reduce the dissimilarity caused by differences in signal amplitude. In this example, the scaling factor is determined for each event pair by grid searching from 0.5 to 2 and retaining the one that maximizes the correlation coefficient.

With the measurements of the absolute travel time and the relative travel-time differences between event pairs, we run the hypoDD program to get refined location results. The double-difference method is appropriate because the local earthquakes occurred in two swarms with diameter less than 2 km according to the locations reported by Raleigh *et al.* (1976). Before the relocation process, we calculate the rough locations using the Velest program (Kissling *et al.*, 1994) with the absolute *P*-wave arrival time alone. The initial location of the hypoDD program is set to be the cluster center of the rough locations. We also set the initial depth to be 2 km corresponding to the shallow cluster in Raleigh *et al.* (1976), because the rough locations are poorly constrained in depth. We run the 20 iterations of the relocation program with increasing weight on the cross-correlation time and decreasing weight on the absolute travel time. We observe that *S* waves dominate the two low-gain channels, so we assume that the correlation measures *S*-wave relative time difference on these two channels.

We correct for the misalignment of different channels caused by image distortion created during either the recording or scanning process. The Inter-Range Instrumentation Group (IRIG) time codes at the top and bottom of the image are identical signals and can be cross correlated with each other to measure the skew of the film. This allows us to correct the skew of the scanned image. The seismic channels in the middle are shifted back assuming linear distortion. The linear skewness distribution is confirmed in the case of electronic noise spikes, which are observed to be constant in time across channels once we correct for skew. Without this correction, the observed time



▲ **Figure 6.** Calculation of the characteristic function for short-term average/long-term average (STA/LTA) picker. (a) Original image slice converted into grayscale. 1 is black and 0 is white. (b) Horizontal gradient of the original image scaled to 0 to 1 for display. 1 is white and 0 is black. (c) $f(t)$: sum of the original image over vertical axes. (d) $f'(t)$: sum of the horizontal gradient image over vertical axes. (e) The characteristic function $E(t)$.

shift errors could be as large as 0.3 s, which would badly degrade the location results.

Unfortunately, the channel listings for the Develocorder recordings were lost. Fortunately, the station locations can be found in a list of the Rangely oil field seismograph stations (Poppe, 1979). Subsequently, we were able to determine the station-trace matches from regional and teleseismic wave arrival order across channels as computed from the catalog locations. Because the seismic waves of the local events have lower apparent velocity across the network, we use some of the clear local events to distinguish the channel order of closely spaced stations around the injection wells.

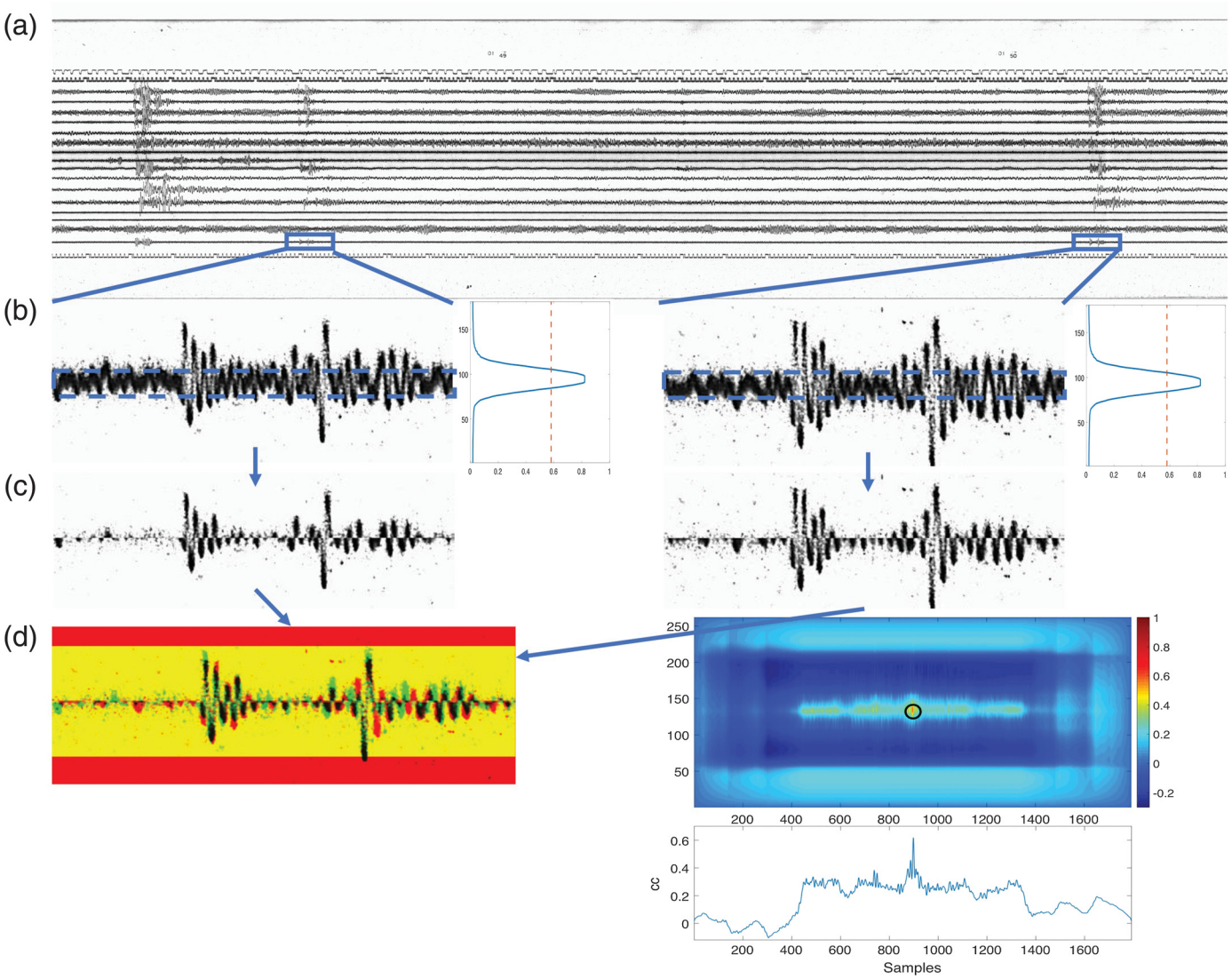
RESULTS

Dataset

We test the performance of our automatic detection and location workflow on the Rangely oil field dataset of scanned Develo corder films. We take one month of the scanned films to demonstrate and validate our processing workflow. Each roll of microfilm records one-day long continuous seismogram. We choose February 1973 as it is in a seismically active period with 32 local events, according to the original report. Although we have processed only one month of data, we scanned the microfilms from April 1972 to February 1974. The remaining films are still being scanned by a ScanStation Lasergraphics film scanner. Under the current settings, it takes 1 hr to scan each day-long film. The detection step also takes about 1 hr without parallelization. The processing time of the location step is mainly on the image correlation between event pairs. It takes approximately 9 hrs to run correlation among the 60 local events on 16 channels.

Detection and Location Results

We demonstrate that our automatic workflow successfully detects and locates more events than the original catalog built by a human analyst. We select 174 earthquake images after the PCA and SVM classification. The STA/LTA picker and the grid-search associator retain 60 local events out of the selected images. We manually check the discarded events and confirm that most of them are regional or teleseismic events. Each of the event pairs in those 60 events is correlated to measure the relative time delays. In the last step, we use both the absolute

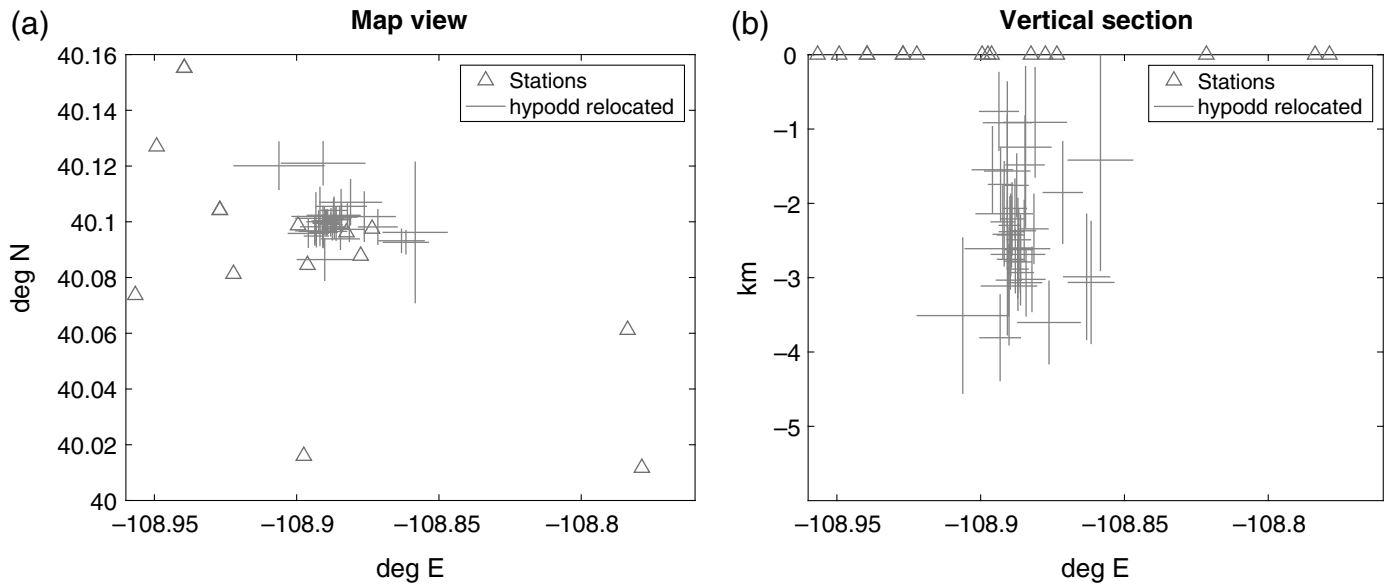


▲ **Figure 7.** Image correlation step. (a) A sample image before processing, (b) selected image windows and the pixel grayscale color distribution to their right, (c) images after removing overlaps and center slices, and (d) correlation coefficient map on the top right and the images overlaid on the best match position on the top left. The bottom right curve shows the normalized correlation coefficient values along the center line.

travel-time picks and the relative time delays in hypoDD to relocate 40 events that form a single cluster below the injection wells. In contrast, the number of events within 1 km of the experimental wells was 32, reported by Raleigh *et al.* (1976). Although the original catalog has been lost, we are confident that we can detect catalog events. It is also reported that the magnitudes of the Rangely earthquakes chosen for the 1976 paper were $M_L > -0.5$, which suggests that our algorithm can detect and locate local earthquakes on the Develocorder films with magnitude down to -0.5.

To validate this correlation-based event relocation and selection step, we manually inspect both the discarded and the relocated events. Four of the 20 discarded events were eliminated by hypoDD, because they come from another cluster and are lost because of lack of connection with the main cluster events. Half of these discarded events are small so that not enough correlation measurements are available for relocation. The rest of the removed events are possibly not similar enough to be connected to the cluster, either because they are not local to this network (~30 km away) or for some other unclear reasons.

Our event locations are shown in Figure 8. The final locations are determined by double-difference relocation program (Waldhauser and Ellsworth, 2000), which better constrains the relative locations of the events. As can be seen in the map view, the events cluster in one swarm and are located within 1 km of the injection wells, which is consistent with the original analysis conducted by a human analyst. The strong consistency in locations may suggest that our location algorithm designed for the Develocorder films could achieve near human performance. Figure 8b shows the location results in vertical section. We compare our locations with those presented in Raleigh *et al.* (1976). Our hypocentral distribution matches the cluster immediately adjacent to the injection wells, at about 2.5 km in depth. The original analysis describes that the epicenters of the earthquakes



▲ **Figure 8.** Final location results with 2-sigma error bars. (a) Map view and (b) vertical section.

are distributed into two clusters with the other one at 3.5 km depth and to the southwest. In our study, however, we see few events from the other cluster. A possible explanation is that the other cluster was inactive during the one-month time period we examined. By manually inspecting the films, we do see four local events with similar waveforms but different from the main cluster. We suspect that these events may come from the deep cluster and might be neglected because there are not enough of them to form a separate cluster in hypoDD. We expect to resolve this issue once we have analyzed data of longer duration.

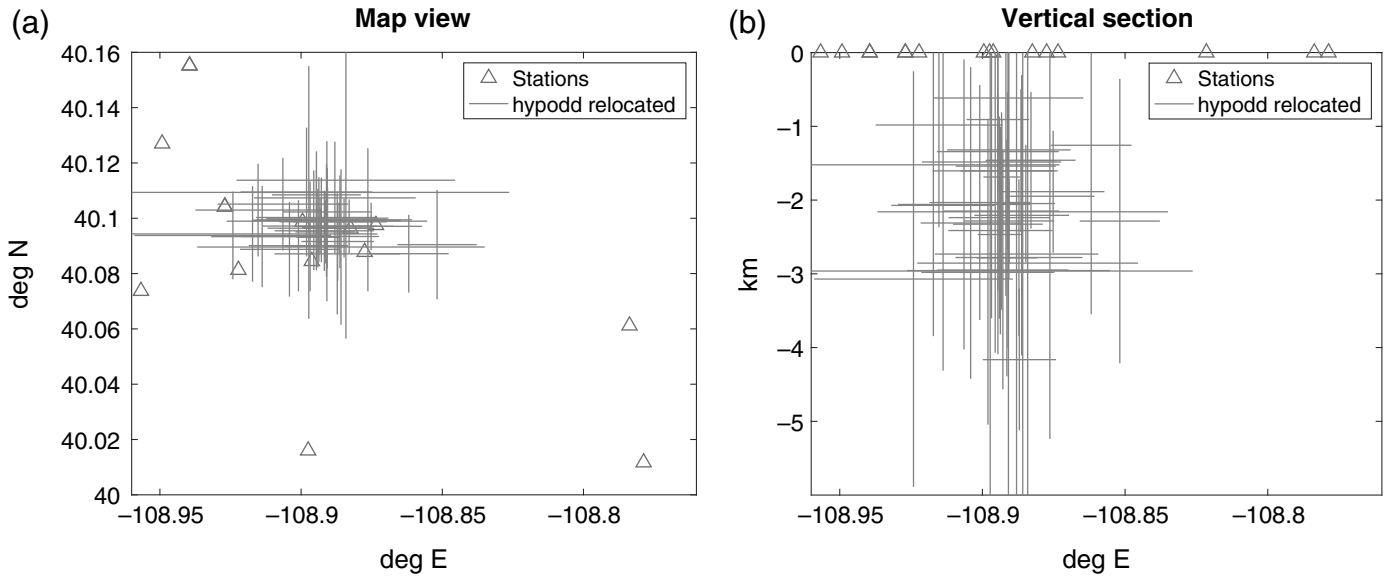
DISCUSSION

Location Uncertainty

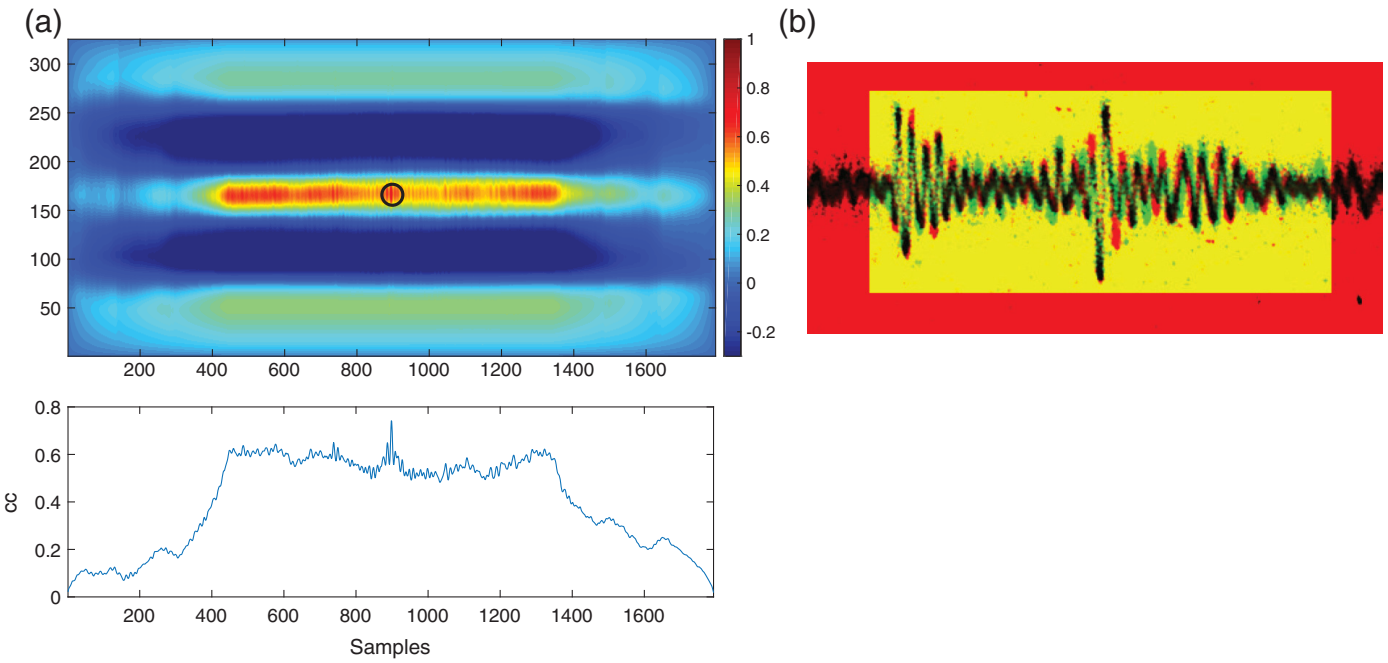
As reported by Raleigh *et al.* (1976), the relative errors in epicentral and depth locations due to different first arrival measurements between two analysts are about 200 and 300 m, respectively. Although this comparison shows the reliability of the manual measurements, it does not necessarily capture the overall uncertainty of the locations. To compare the uncertainty of our location results with the original analysis, we manually pick the *P*-wave arrival time on the films and follow the previous steps to get initial locations using the Velest program and refine the relative locations using hypoDD. The original analysis was conducted using the HYPOLAYR program, a least-squares hypocenter location program for a layered model. The program makes limited use of *S*-phase data (Eaton, 1970) by computing origin time with *S*–*P* measurements and not using the *S*-wave arrival times themselves to adjust the hypocenter. We believe that using the double-difference location algorithm will reduce the uncertainty for relative locations compared with that of the original study. The comparison of the location results between manual picking and the automatic workflow is shown in Figure 9. As can be seen in Figure 9, the location uncertainties for the manual processing are larger than those from the automated workflow. This advantage is possibly attributable to our use of cross-correlation time delay measurements. As in other cases, we would expect this advantage to increase for swarms with more events, which could have more correlated event pairs.

Potential to Exceed Human Level Performance

Unlike modern digital seismograms, the Develocorder film recordings are made by exposing photographic film with a light beam that have finite and fixed intensity. As the beam moves quickly, the recorded intensity fades and the spot shrinks in image. Figure 10 shows the correlation map on the left corresponding to the overlain image of two earthquakes on the right. The circle indicates the position of maximum correlation used to measure a relative time delay. In Figure 10b, the yellow area corresponds to the area we correlate, black pixels are common to both images, and the red and green pixels are unique to one or the other event. From the correlation maps, we observe higher correlation coefficients along the trace than in Figure 7d, which suggests correlation of noise pixels make a non-negligible contribution to the correlation coefficients. Therefore, we mask the center part of the image in the correlation step to avoid correlating noise. Based on our experiences, it may be possible to measure time delay to less than the beam width; however, we still have the problem of losing information when the trace fades, which is especially true for stronger events. It should be possible to improve this situation by considering the time when the trace center has an abrupt change in color density. Another point is that our detection model is only trained on local earthquakes. As a result, it is not sensitive to regional or teleseismic events. Separate models need to be trained for detecting those events. Our current processing workflow works reasonably well compared with the results of the original study, which was based on readings of a single human analyst. With these modern image processing approaches, we expect in future versions we could



▲ **Figure 9.** Location results by manual picking with 2-sigma error bars. (a) Map view and (b) vertical section.



▲ **Figure 10.** (a) The correlation map of two events on the top. The circle indicates the maximum correlation coefficient position. The bottom shows the normalized correlation coefficients along the center line. (b) The two events overlaid at the time delay corresponding to the maximum correlation coefficient.

achieve much higher timing accuracy than possible with a ruler on the Geotech viewer.

CONCLUSIONS

In this article, we demonstrate a new workflow to detect and locate local events from Develocorder films. To do this, we scan the films into images and directly apply image-based methods without converting the waveform images into time series. Our method detects more events than the original catalog. After detection, we locate those events using both absolute arrival times and relative time delays between event pairs. The resulting locations have lower uncertainties than those obtained from the manually picked first arrivals.

A distinct advantage over other approaches to working with analog data is that we do not need to convert the data to a time series. We developed our method using Develocorder films, but an image-based detection and location

approach should be easily adaptable to analog data on diverse recording media from the predigital era.

The Rangely case provides a good example of the motivation to work with older data. Half a century has passed and much has been learned about earthquake physics since the Rangely experiment. Once we study these data comprehensively, we can make full use of this unique human-controlled injection experiment to test the predictions of rate-and-state theory (Dieterich, 1979, 1994; McClure and Horne, 2011), for one, or compare predictions of full poroelastic coupling of stress and pore pressure (Segall and Lu, 2015) that were not yet formulated at the time of the experiment.

DATA AND RESOURCES

The Develocorder films are provided by U.S. Geological Survey (USGS). The following catalogs are used: USGS catalog (https://earthquake.usgs.gov/earthquakes/), Utah Region catalog July 1962–September 1974, and Yellowstone USGS Catalog 1973–1981 from the University of Utah Seismograph Stations (both from http://quake.utah.edu/regional-info/earthquake-catalogs). All websites were last accessed August 2018.

ACKNOWLEDGMENTS

This work was carried out with support from the Stanford Center for Induced and Triggered Seismicity (SCITS). The authors thank Ole Kaven and Morgan Page for reviews in preparation of the article.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

REFERENCES

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68,** 1521–1532.

Baskoutas, I., I. Kalogeras, M. Kourouzidis, and G. Panopoulou (2000). A modern technique for the retrieval and processing of historical seismograms in Greece, *Nat. Hazards* **21,** 55–64.

Benson, R., W. H. K. Lee, C. R. Hutt, T. Knight, T. Ahern, M. F. Diggles, and S. R. Walter (2011). SeismoArchives online at the IRIS DMC; preserving pre-digital seismograms (1882–1990) for research & information [abs.], *2011 International Union of Geodesy and Geophysics (IUGG) General Assembly*, Melbourne, Australia, 28 June–7 July.

Bogiatzis, P., and M. Ishii (2016). DigitSeis: A new digitization software for analog seismograms, *Seismol. Res. Lett.* **87,** 726–736.

Cortes, C., and V. Vapnik (1995). Support-vector networks, *Mach. Learn.* **20,** 273–297.

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.* **99,** 2601–2618.

Dieterich, J. H. (1979). Modeling of rock friction: 1. Experimental results and constitutive equations, *J. Geophys. Res.* **84,** 2161–2168.

Eaton, J. P. (1970). HYPOLAYR, A computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half space, *U.S. Geol. Surv. Open-File Rept. 69-85*, doi: 10.3133/ofr6985.

Gibbs, J. F., J. H. Healy, C. B. Raleigh, and J. M. Coakley (1972). Earthquakes in the oil field at Rangely, Colorado, *U.S. Geol. Surv. Open-File Rept. 72-130*, available at http://pubs.usgs.gov/of/1972/0130/ (last accessed November 2018).

Ishii, M., H. Ishii, B. Bernier, and E. Bulat (2014). Efforts to recover and digitize analog seismograms from Harvard-Adam Dziewoński observatory, *Seismol. Res. Lett.* **86,** 255–261.

Jolliffe, I. (2011). Principal component analysis, in *International Encyclopedia of Statistical Science*, Springer, Berlin, Germany, 1094–1096.

Kissling, E., W. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer (1994). Initial reference models in local earthquake tomography, *J. Geophys. Res.* **99,** 19,635–19,646.

McClure, M. W., and R. N. Horne (2011). Investigation of injection-induced seismicity using a coupled fluid flow and rate/state friction model, *Geophysics* **76,** WC181–WC198.

O'Neill, M., and J. Healy (1973). Determination of source parameters of small earthquakes from *P*-wave rise time, *Bull. Seismol. Soc. Am.* **63,** 599–614.

Pintore, S., M. Quintiliani, and D. Franceschi (2005). Teseo: A vectoriser of historical seismograms, *Comput. Geosci.* **31,** 1277–1285.

Poppe, B. B. (1979).Historical survey of US seismograph stations, *U.S. Geol. Surv. Profess. Pap. 1096*, 398 pp.

Raleigh, C. B., J. H. Healy, and J. D. Bredehoeft (1976). An experiment in earthquake control at Rangely, Colorado, *Science* **191,** 1230–1237.

Segall, P., and S. Lu (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects, *J. Geophys. Res.* **120,** 5082–5103.

Turk, M., and A. Pentland (1991). Eigenfaces for recognition, *J. Cogn. Neurosci.* **3,** 71–86.

Waldhauser, F. (2001). hypoDD—A program to compute double-difference hypocenter locations (hypoDD version 1.0-03/2001), *U.S. Geol. Surv. Open-File Rep. 02-1123*, doi: 10.3133/ofr01113.

Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California, *Bull. Seismol. Soc. Am.* **90,** 1353–1368.

Wang, M., Q. Jiang, Q. Liu, and M. Huang (2016). A new program on digitizing analog seismograms, *Comput. Geosci.* **93,** 70–76.

Xu, Y., and T. Xu (2014). An interactive program on digitizing historical seismograms, *Comput. Geosci.* **63,** 88–95.

Kaiwen Wang
William L. Ellsworth
Gregory C. Beroza
Miao Zhang
Dustin Schroeder
Department of Geophysics
Stanford University
397 Panama Mall
Stanford, California 94305 U.S.A.
kaiwenw@stanford.edu

Gordon Williams
Department of Earth Sciences
University of California Santa Barbara
Lagoon Road
Santa Barbara, California 93106 U.S.A.

Justin Rubinstein
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025 U.S.A.

Published Online 12 December 2018