

УДК 550.34

СОЗДАНИЕ ЦИФРОВОГО АРХИВА БУМАЖНЫХ СЕЙСМОГРАММ В КФ ФИЦ ЕГС РАН: ПЕРВЫЕ РЕЗУЛЬТАТЫ

Ромашева Е.И., Матвеев Е.А., Чеброва А.Ю., Шевченко С.А.

*Камчатский филиал ФИЦ «Единая геофизическая служба РАН», г. Петропавловск-Камчатский,
evarom@emsd.ru*

Введение

В Камчатском филиале ФИЦ ЕГС РАН ведется разработка Единой информационной системы сейсмологических данных (ЕИССД), которая обеспечивает сбор, систематизацию, контроль, хранение и доступ к сейсмологической информации в едином информационном пространстве. ЕИССД постоянно развивается, объем информации, содержащейся в базе данных (БД) ЕИССД непрерывно растет [1]. Система состоит из ряда подсистем, обеспечивающих выполнение различных задач в рамках сейсмического мониторинга.

В 2019 г. в КФ ФИЦ ЕГС РАН началась работа по сканированию бумажных сейсмограмм, записанных аналоговыми стационарными станциями Камчатской сети, работавшими в период 1962–2006 гг. Эта работа осуществляется в рамках проекта развития ЕИССД и преследует несколько целей:

1) Создание архива цифровых копий бумажных сейсмограмм за период 1962–2006 гг., дополняющего уже имеющийся в КФ ФИЦ ЕГС РАН архив непрерывных сейсмических данных цифровых сейсмических станций за период 1993–наст. вр.

2) Сбор метрик работоспособности сейсмометрических каналов (информация о датах начала и окончания записей сейсмических станций, о пропусках в записях сейсмометрических каналов и др.) путем загрузки метаданных о бумажных сейсмограммах в БД ЕИССД.

3) Дополнение имеющегося в КФ ФИЦ ЕГС РАН архива волновых форм [2, 3] землетрясений за период 1996–наст. вр. путем преобразования изображений записей землетрясений в цифровые сейсмические трассы (оцифровка записей землетрясений).

В работе ведется речь о методике создания цифрового архива бумажных сейсмограмм, включающего два этапа: 1) сканирование сейсмограмм и 2) сохранение цифровых копий (сканов) сейсмограмм и их метаданных в БД ЕИССД с помощью специально созданного инструмента. Кроме того, рассказывается о первых попытках в вопросе оцифровки записей землетрясений.

Методика создания цифрового архива бумажных сейсмограмм

Сканирование сейсмограмм. В архиве КФ ФИЦ ЕГС РАН содержится более 1 млн. бумажных сейсмограмм, начиная с 1962 по 2006 гг. Не смотря на соблюдение рекомендаций по условиям хранения данных, ленты с непрерывными сейсмическими записями начинают желтеть и крошиться.

Бумажная сейсмограмма представляет собой лист фотобумаги размера 91.5 на 28.5 см. На одной стороне находится непрерывная запись сейсмических каналов: север-юг, запад-восток, верх-низ, заглубленный (не всегда присутствует) (рис. 2а). На обратной стороне сейсмограммы имеется штамп, в который внесены от руки метаданные о сейсмограмме: название станции, порядковый номер ленты в году, дата и время записи, поправки по времени, увеличение на канале, подпись оператора сейсмостанции (рис.2б).



Рис. 1. Фото помещения архива КФ ФИЦ ЕГС РАН

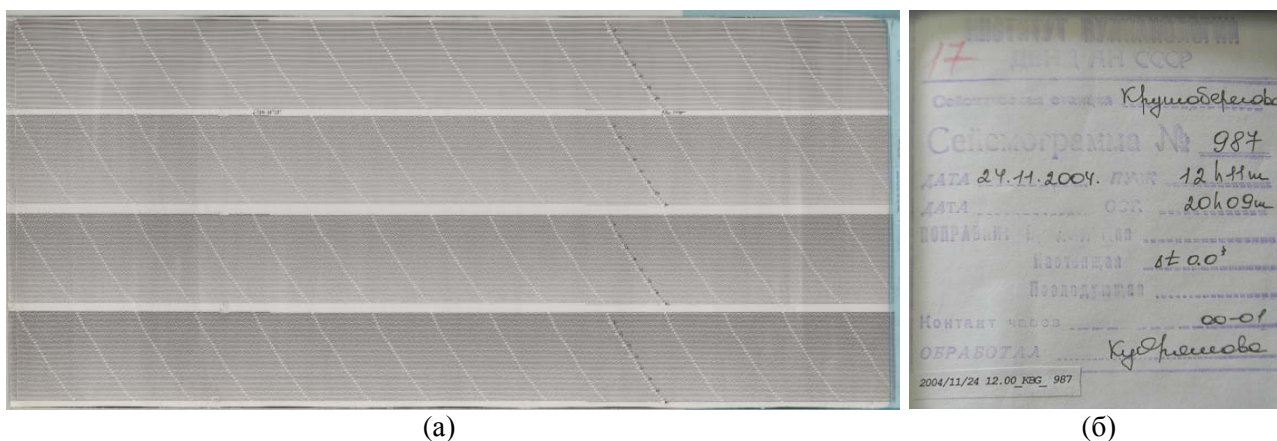


Рис.2. Восьмичасовая непрерывная сейсмическая запись на фотобумаге (а) и штамп с обратной стороны сейсмограммы (б).

Для осуществления оцифровки сейсмограмм был приобретен широкоформатный протяжный сканер (рис. 3), максимальная ширина сканирования 91.4 см. Этого размера хватает впритык, в некоторых случаях приходится отрезать несколько мм с края ленты, не имеющего информации.

При обработке мы столкнулись с ветхостью и заломами некоторых сейсмограмм. Поставщиком сканера был предложен вариант приобретения специального защитного файла для сканирования ветхих бумажных носителей. Но в тот момент у нас не было возможности приобрести этот достаточно дорогой аксессуар. Силами сотрудников был найден выход из положения – использование пластикового уголка для «легкого» прохождения бумажных сейсмограмм через сканер. Сканирование ведется в цвете с разрешением 600 dpi, тип сохраняемого файла PNG, средний размер полученной отсканированной сейсмограммы составляет 200–210 МБ.

Штамп на оборотной стороне сейсмограммы не сканируется, а фотографируется. Для этого была специально сконструирована рамка по размерам штампа. Это позволяет не тратить время на обработку (обрезку) изображения и экономить место на сервере, средний размер фотографии штампа составляет 1–1.5МБ.

Отсканированные сейсмограммы и фотографии штампов сохраняются в структурированные папки (...\\Год\\Порядковый_номер_сейсмограммы_в_году). Следует отметить, что порядковый номер сейсмограммы в году соответствует определенному времени записи.

На сегодняшний день отсканировано более 3100 сейсмограмм, что составляет менее 1% архива. Работа ведется параллельно с другими задачами лаборатории в промежутки времени, свободные от более срочных и важных поручений.



Рис.3. Широкоформатный сканер, приобретенный для сканирования сейсмограмм.

Сохранение цифровой информации в ЕИССД. Все отсканированные сейсмограммы и сопровождающая их вспомогательная информация (метаданные) сохраняется в архив и БД ЕИССД соответственно. Для этих целей был разработан веб-интерфейс для ввода и сохранения данных об отсканированных сейсмограммах (рис. 4).

Веб-страница разделена на три части (слева – поля для ввода метаданных сейсмограммы, справа – привью ленты (уменьшенная до небольших размеров копия изображения), в центре – фото печати с обратной стороны ленты).

На обе стороны бумажного носителя до сканирования наклеиваются стикеры, содержащие шифр сейсмограммы в формате Год/Месяц/День_Время_КодСтанции_НомерСейсмограммы. При внесении информации оператору необходимо удостовериться, что шифры на печати и ленте совпадают, после этого информация, которая содержится на печати, вводится в соответствующие поля в левой части страницы.

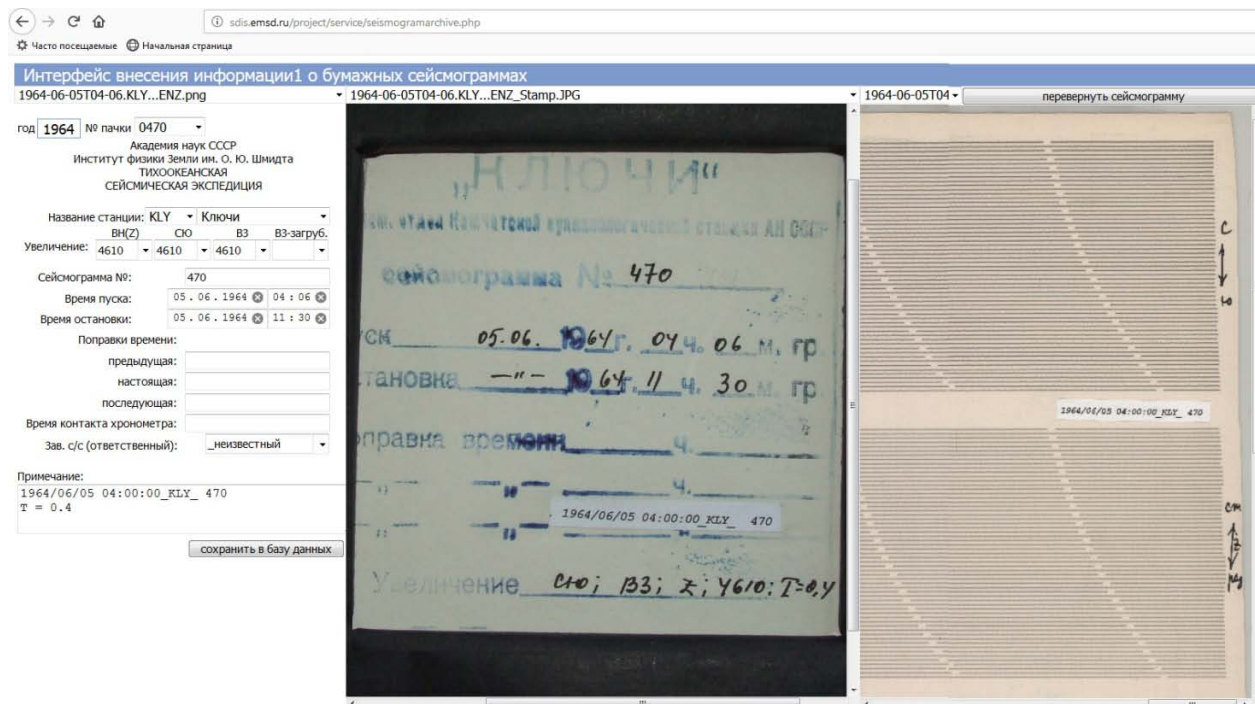


Рис.4. WEB-интерфейс для ввода метаданных с отсканированных сейсмограмм.

Следует отметить, что большая часть информации подгружается в поля автоматически:

- ✓ после выбора года и номера сейсмограммы автоматически заполняются дата и время;
- ✓ на основании названия станции, автоматически подгружаются станционные данные, содержащиеся в базе данных (увеличение на каналах, поправки по времени, время контакта хронометра, ответственный оператор станции);
- ✓ шифр сейсмограммы со стикера автоматически добавляется в поле «Примечание».

Полей для ручного заполнения остается не много. Поля, заполненные автоматически, необходимо проверить, а отсутствующие или подгруженные неверно внести в соответствии с данными на печати.

После внесения информации печать, привью ленты и исходный файл ленты сохраняются на сервер в архив, метаданные и информация о месте хранения исходного файла сохраняются в БД ЕИССД.

Оцифровка аналоговых записей землетрясений

Параллельно с переносом бумажных носителей в электронный вид и наполнением ЕИССД мы пытались начать работу по оцифровке землетрясений. Нами была найдена бесплатная программа DigitSeis сейсмологической группы Гарвардского Университета. Данная программа является пилотной, она разрабатывается совместно с участием старшеклассников из Японии, конечной ее целью является получение полноценного научного проекта.

DigitSeis представляет собой программный продукт, написанный в Matlab, состоящий из кусочков последовательно выполняемых подпрограмм.

При запуске программного продукта нам предоставляется выбор загрузить новую сейсмограмму или использовать ранее сохраненный процесс с результатами анализа сейсмограммы. При первичной загрузке сейсмограммы программа запрашивает разрешение на конвертацию изображения, так как она распознает только светлые трассы на темном фоне. Так же в этот момент идет преобразование цветного изображения в градации серого. Далее мы указываем размер шума, размер временных меток. В процессе удаления шума, мы можем производить корректировку: можем поменять размер шума или указать, что конкретно на сейсмограмме является шумом, а что временной меткой или непосредственно трассой. Есть возможность выделения и удаления сразу большого объема ненужной нам информации с сейсмограммы, что позволяет не тратить время на «очистку» изображения перед обработкой в сторонних программах. После обработки изображения начинается этап автоматического распознавания трассы, далее автоматически временным меткам присваиваются значения, на основании ранее заданной информации, и можно выгрузить полученные данные в формате SAC. К сожалению, программный продукт DigitSeis еще находится в стадии разработки и при оцифровке одной и той же сейсмограммы может выдавать разные конечные данные, нельзя вернуться к предыдущему шагу и что-то исправить. Поэтому на данный момент нам не удалось добиться стабильности при ее использовании, хотя понравились идея и реализация.

Вторым протестированным нами программным продуктом была программа, разработанная Новосибирским региональным центром геоинформационных технологий в 1999–2000 гг. – WaveTrack.

WaveTrack позволяет распознавать сейсмические события в автоматизированном режиме: либо автоматически, предварительно очищенные от фона сейсмограммы; либо вручную, расставляя точки на изображении сейсмограммы на экране монитора. Мы протестировали ручное распознавание. Установив время начала сейсмограммы и задав сетку по ранее известным нам значениям, нами были получены цифровые сейсмограммы в виде текстового файла в формате ASCII. Достоинствами программы WaveTrack для нас стали стабильность работы программы, простота и понятность использования.

После выбора подходящей программы для оцифровки записей землетрясений планируется оцифровать некоторую часть сейсмограмм, включающую наиболее интересные события камчатского региона.

Заключение

Сканирование архивных данных – работа объемная, рассчитанная на длительный срок. Пока мы находимся только в начале этого долгого пути. Наша задача сохранить в ЕИССД всю информацию, которую содержит уникальный архив непрерывных записей аналоговой камчатской сейсмической сети за период 1962–2006 гг. А также создать удобные сервисы, позволяющие обратиться как к первоисточнику информации в случае пересмотра, дополнения или уточнения данных обработки, так и к его цифровым производным.

Список литературы

1. *Чеброва А.Ю., Чебров В.Н., Матвеев Е.А., Токарев А.В., Чемарёв А.С.* Единая информационная система сейсмологических данных в Камчатском филиале Геофизической службы РАН по состоянию на середину 2015 года. Современные методы обработки и интерпретации сейсмологических данных / Материалы Десятой Международной сейсмологической школы Азербайджан, 14-18 сентября 2015, Обнинск. С. 356–360.
2. *Бахтиярова Г.М.* Цифровой архив региональных станций Камчатского филиала ГС РАН. // Геофизический мониторинг Камчатки. Материалы научно-технической конференции 17-18 января 2006 г. Петропавловск-Камчатский: КФ ГС РАН, 2006. С. 29–31.
3. *Чемарёв А.С., Токарев А.В.* Подсистема ЕИС для формирования и хранения волновых форм землетрясений // Проблемы комплексного геофизического мониторинга Дальнего Востока России. Труды Четвертой научно-технической конференции. Петропавловск-Камчатский. 29 сентября – 5 октября 2013 г. Обнинск: ГС РАН, 2013. С. 359–361.