SUPRA: DIGITIZING THE
STANFORD UNIVERSITY PIANO ROLL ARCHIVE

Zhengshan Shi1 Craig Stuart Sapp2 Kumaran Arul3 Jerry McBride4 Julius O. Smith III1
1 Center for Computer Research in Music and Acoustics, Stanford University
2 Center for Computer Assisted Research in the Humanities, Stanford University
3 Department of Music, Stanford University
4 Music Library and Archive of Recorded Sound, Stanford University
kittyshi@ccrma.stanford.edu, craig@ccrma.stanford.edu

ABSTRACT
This paper describes the digitization process of a large collection of historical piano roll recordings held in the Stanford University Piano Roll Archive (SUPRA), which has resulted in an initial dataset of 478 performances of pianists from the early twentieth century transcribed to MIDI format. The process includes scanning paper rolls, digitizing the hole punches, and translating the pneumatic expression codings into MIDI format to create expressive performance files. We offer derivative files from each step of this process, including a high resolution image of the roll, a “raw” MIDI file of hole data, an “expressive” MIDI file that translates hole data into dynamics, and an audio file rendering of the expressive MIDI file on a digital piano sample. This provides digital access to the rolls for researchers in a flexible, searchable online database. We currently offer an initial dataset, “SUPRA-RW” from a selection of “red Welte”-type rolls in the SUPRA. This dataset provides roll scans and MIDI transcriptions of important historical piano performances, many being made available widely for the first time.

1. INTRODUCTION
1.1 Background
Piano rolls are among the most important historical music storage formats, utilizing holes on a scrolling paper roll to activate keys automatically by a pneumatic mechanism built into a piano. There are many types of piano rolls with varying degrees of autonomous playback prescribed by the roll system. We focus here on the “reproducing roll”, one that reproduces automatically notes, timings, and expressive details of a live performance.

By late 1904 the German company Michael Welte and Söhne developed the first “reproducing player piano”, the Welte-Mignon [7]. This was a high-end, fully autonomous self-playing instrument capable of reproducing all features of an original performance of a pianist, including details of timing, dynamics and pedaling. The process involved capturing and coding expressive details into the hole punches on the edges of rolls which then cued the pneumatic mechanisms in the piano to alter hammer velocities instantaneously. These codings were created by skilled technicians who made judgements about the expressive effects they could generate using the pneumatic parameters available for manipulation in each system. The resultant playback of a reproducing roll sounds like a recreation of an individual musical performance.

The reproducing player piano quickly became a commercial success that attracted the most important artists of the day to make recordings on the medium. Composers recorded playing their own works on roll include Claude Debussy, Maurice Ravel, Gustav Mahler, Sergie Prokofiev, Sergei Rachmaninoff, Edvard Grieg, Alexander Scriabin, Enrique Granados, Igor Stravinsky, George Gershwin, and Scott Joplin, among others.

A variety of proprietary reproducing roll systems were produced by over a dozen companies, each one incompatible with the others. The most important of these include Welte-Mignon, Hupfeld, Ampico, Duo-Art, and Philips Duca. Some companies developed multiple formats. Welte manufactured three formats: Welte T-100 (red Welte), Welte T-98 (green Welte), and Welte-Licensee.

1.2 SUPRA
Stanford University’s Player Piano Project [13] is a multi-departmental effort begun in 2014 by the Archive of Recorded Sound to address the obstacles faced by researchers who wish to study piano rolls and pneumatic instruments. Stanford now holds one of the largest roll collections in the world, with more than 16,000 rolls and a dozen pneumatic roll playing instruments. The collection includes a variety of roll types, including thousands of reproducing rolls. A dedicated roll scanner was built by the project to generate image scans of rolls for preservation and digitization. The Stanford University Piano Roll Archive (SUPRA) is the online database of roll images generated by the scanning effort. It aims to provide a virtual experience of piano rolls including a high resolution
color image of the roll and a digitally synthesized audio rendition of each roll. In order to create accurate audio files of reproducing rolls, algorithms that model the pneumatic expression systems of each format must be created. This process of “emulation” (i.e., emulating the pneumatic system through a digital means) is important if the expressive content of reproducing rolls are to be accurately transcribed. This paper outlines the end-to-end process undertaken to scan, digitize, and create emulation algorithms for a subset of 478 Welte T-100 reproducing rolls that will be made accessible through the online SUPRA database.

2. PRIOR WORK

Efforts at scanning and digitizing piano rolls have been undertaken by hobbyists and enthusiasts with private collections, but with inconsistent results and incomplete documentation [1, 15, 18]. More recently a number of institutional projects have been initiated [2, 3, 8–10]. Much of this work has focused on non-reproducing rolls, those rolls lacking expression coding [8, 12]. Most institutional efforts do not provide audio file transfers of the rolls at this point [2, 3, 9]. In some cases only documentation of the rolls is offered but not complete roll scans [3].

A number of dedicated roll scanners have been constructed [1, 9, 12, 15, 18]. However, some projects utilize flat bed scanners which require images to be stitched together leading to potential errors [8]. Peter Phillips’s unique design of a “pneumatic roll reader” offers some advantages, but does not allow for an archival image [14]. Anthony Robinson has done impressive work with scanner design and is able to produce quality roll scans [15]. These are, however, in gray scale and at modest resolution.

There is not much published work on the topic of emulation although there have been some well publicized efforts at utilizing emulations to playback reproducing rolls. Colmnares et al. [6] provide some theory behind the topic and relies on the work of Wayne Stahnke, one of the pioneers in modeling pneumatic roll systems. Stahnke produced two commercially successful emulated transfers of Rachmaninoff’s playing on Ampico reproducing rolls that garnered critical acclaim [17], however the algorithms and detailed procedures of his work remain unpublished. Peter Phillips’s recent doctoral thesis provides comprehensive documentation and experimental justification for his emulation procedures and is the most thorough discussion published thus far [14]. Our work here follows closely on his effort and also on our prior work to create an emulation for the Welte Licensee format [16].

3. THE SUPRA-RW DATASET

We chose to begin the SUPRA database with Welte T-100 rolls as these were the first reproducing piano rolls made. These important recordings have received very little attention from roll scanning efforts because they require a scanner capable of managing the significantly greater width of these rolls. Expression emulation of Welte T-100 rolls is also among the most complex of reproducing roll formats.

Figure 1: A sample from a red Welte roll. It is divided into four sections: bass expression, bass notes, treble notes, and treble expression.

Welte T-100 rolls are often referred to as “red Welte rolls” because they were generally punched on red paper. They are 12.9 inches (32.8 cm) wide, and have 100 perforation tracks (holes), eight per inch across, evenly spaced. A sample from a red Welte roll is shown in Figure 1. The first and last ten perforation tracks are used for expression, and the middle 80 tracks are used for notes from C1 to G7.

The SUPRA-RW dataset is the result of the digitization of 478 Welte T-100 rolls in Stanford’s collection. It consists of about 52 hours of piano roll performances, with an average length of 6 minutes and 32 seconds. The database allows users to search, view images, download and listen to audio emulations of the rolls.

Our digitization procedure results in the following files and derivatives in sequence:

(a) Archival TIFF image at 300 DPI (dots per inch) and 24-bit color.
(b) JPEG files derived from the archival TIFF.
(c) Uncompressed grayscale TIFF file (that utilizes the green color channel of the original scan) allowing efficient access to high resolution detail of the holes.
(d) Raw MIDI file that captures all hole data extracted from the grayscale TIFF file.
(e) Expressive MIDI file that merges multiple holes into single musical notes, applies emulation algorithms to control for individual note dynamics, and adds pedaling. Metadata such as title and composer is also added to this file.
(f) Audio files (WAV and M4A) rendered by running the expressive MIDI file through the Ivory Keys II software synthesizer.

The SUPRA-RW dataset as well as the software are available online\(^1\) at a Creative Commons Attribution-Non-Commercial-ShareAlike 4.0 International license (CC BY-NC-SA 4.0)\(^2\).

---

\(^1\) https://supra.stanford.edu
\(^2\) https://creativecommons.org/licenses/by-nc-sa/4.0/
4. ROLL DIGITIZATION

4.1 Scanning

Piano rolls are digitized using a dedicated roll scanner built in conjunction with Swope Design Solutions and based on designs by Anthony Robinson\(^3\). The scanner (Figure 2) uses a line-scan camera (DALSA Spyder3 Color 4k) triggered by a rotary encoder on a glass cylinder that tracks the movement of the paper roll that passes over it. The target resolution of the images is 300 DPI, with the experimentally measured DPI being 301.50 ± 0.25 across the width of a roll and 300.25 ± 0.25 along the length of the roll.

The scanner can image a roll up to 5 × playback speed, although the typical acquisition speed is 2–3×. An original uncompressed color image is typically 1–4 GB in size. The line-scan camera takes pictures that are 2 pixels high and 4096 pixels wide. One row of pixels is used to measure the green color channel, and the second row contains interleaved red and blue color pixels. The adjacent images from the camera are 50% overlapped such that a red/blue row from one image is aligned with the green row of the next image. Since the green channel has twice the resolution of the red or blue channels, the green channel is used for hole data extraction from the images.

\(^3\) For more details and a video of the scanner in action, see https://library.stanford.edu/blogs/stanford-libraries-blog/2018/10/piano-roll-scanner-update.

![Figure 2: Stanford’s piano roll scanner built by Swope Design Solutions of San Francisco, California. A Welte T-100 piano roll is shown on the scanner.](image)

4.2 Drift Correction and Hole Detection

The paper of an original piano roll can warp due to age and poor storage conditions. In such cases, paper holes may drift and misalign with the corresponding tracker bar hole and trigger the incorrect note. Figure 3 shows the spacing between the rectangular tracker bar holes and the circular holes of the piano roll paper. When the paper shifts by more than one mm (about 13 pixels), a paper hole will start to bleed the vacuum in an adjacent tracker bar hole. If the overlap is large enough, an incorrect note will be triggered on the player piano. Thus, a first step in extracting the musical holes from a roll image involves identifying and cancelling this drift.

Figure 4 plots the drift of a sample roll throughout its length. There is an initial sudden shift of about 15 pixels at the start of the roll caused by the scanning operator moving an adjustment bar. For the next 25 feet the drift is minimal, at only a few pixels. Then starting at around 27 feet, an oscillation with an amplitude of 20 pixels begins to be observed. This oscillation has a frequency of about 5–6 feet and is common in many Welte T-100 scans. We determined this to be due to inconsistencies in the paper, as well as mechanical properties of the scanner relating to the tension and alignment of the supply and take-up spools.

For a sample of 60 red rolls, the average total drift range was 20.8 ± 10.8 pixels, where the spacing between note columns was 37.75 px. Five of the rolls had a total drift that would cause them to produce incorrect notes if played without correction on a player piano, and 25 had a drift that may cause problems. The largest total drift throughout the
scan of a roll was 55.7 px, and the smallest was 2.7 px over the length of a roll.

To calculate the drift correction along the length of a piano roll, the edges of the paper are identified, and these edges are low-pass filtered to remove defects such as edge tears from the edge analysis. When the left and right edges of the paper move suddenly in parallel, this indicates that the scanner operator moved the adjustment bar during the scanning process.

This left-right drift is cancelled out by adjusting the horizontal position by the negated drift value at that position on the roll. The centroid (center-of-mass) for each musical hole is identified in the image, and then assigned a corrected position by subtracting the drift offset at the centroid position. Figure 5 shows the results of this drift correction for a single tracker bar position. After adjusting for drift, the hole centroids for a track are clustered within a standard deviation of less than one pixel (1/300th of an inch), compared to a range of 18 pixels before the adjustment.

Finally, a Discrete Fourier Transform (DFT) is applied to the drift-corrected hole centroid histogram (window size 4096, zero padding by a factor of 16). The spacing of the hole tracks is obtained by searching for the peak harmonic in the expected region of the DFT. The offset of the spacings is locked to the track position with the most hole punches.

4.3 Defect Analysis

The drift adjustment also allows for increased accuracy in identifying holes that are not quite aligned with any intended tracker hole. Such errors can be removed in this step. Figure 6 illustrates such a problematic hole. In this example, the green-highlighted hole at the bottom left corner of the figure is caused by a tear in the original paper. This hole is not aligned on any of the expected hole centers that are indicated by purple vertical lines. The blue holes indicate alignment as expected.

Other statistical measures of hole shapes are also used to detect aberrant holes. For example, the angle of the major axis of the hole is expected to align with the length of a roll, except for circular single-punch holes. Holes less than 1/6 of the expected area of a single punch are automatically excluded. These are typically small defects in the paper, such as dirt or large pieces of cellulose in the paper that have fallen out over time. Holes wider than the expected spacing between holes are also flagged as potential problems.

4.4 Bridge Removal

Long holes on piano rolls are often split into several smaller holes to avoid weakening the paper, however they sound as a single longer note when played back on a pneumatic instrument. This process is called “bridging”. Our raw MIDI files retain the bridging information, thus making them suitable for punching new paper copies of the rolls (what are called “recuts”) that retain these separations. This detail is also important for master-roll reconstruction analysis. Our expressive MIDI files however remove this bridging to produce the resultant sound of the roll. Adjacent holes are taken to merge into single notes when the spacing between holes is less than 1.37 × the diameter of the punches. Figure 6 illustrates such a case where the two blue holes in the upper right corner represent a single note and would be merged.

5. ROLL EMULATION

5.1 The Pneumatic System

A player piano is powered by suction, utilizing pneumatic valves that regulate vacuum pressure created by an electric motor. As shown in Figure 7, when a roll passes over the tracker bar, a hole on the paper allows air to leak into the lower section of the valve box, causing the air below to change to atmospheric air pressure. As it moves the valve pouch upwards, the vacuum thus travels to the pneumatic and collapses it, causing the pneumatic to move, which strikes a piano key. However, on a reproducing piano, a further series of pneumatic controls, the “expression box”, regulates the precise rate of change in the suction level of the “dynamic” pneumatic. There is one expression box for each half of the keyboard, thus allowing dynamic changes to only occur on one half at a time, however the changes
can occur very rapidly over time. Thus, the dynamic pneumatic can do one of the following:

(a) Remain stationary.
(b) Open or close slowly, producing the effect of a slow crescendo or decrescendo (suction level increases or decreases slowly).
(c) Open or close quickly, producing the effect of a fast crescendo or decrescendo (suction level increases or decreases quickly).

5.2 Expression Emulation of Welte T-100 Rolls

The process of emulation involves understanding how the expression box (or equivalent component in other formats) affects suction pressure on the dynamic pneumatics for each key over time. The expression system for Welte T-100 rolls involves ten parameters (including pedals) that can affect independently or in combination the dynamic result at any given time. Table 1 shows the possible parameters and their locations as expression hole tracks. There are parallel holes for each expression regulation on each side of the piano because, as mentioned above, the pneumatic expression mechanism is split in half across the keyboard. The damper (sustain) pedal plays a role in the dynamics because it affects the collective sustain and decay of notes played.

We begin by encoding each hole track as a MIDI note number in the raw MIDI file. The function of each hole track on red Welte rolls is summarized in Table 1. Holes of the bass expression tracks (MIDI Note Number 14—19) control the dynamics of notes below F#4 and those of the treble expression tracks (108—113) control the dynamics of notes above G4. Holes for pedal movements (20—21, and 106—107) apply to all notes. Notes (pitches) on the keyboard are holes 24–103.

The “Mezzoforte on” and “Mezzoforte off” hole tracks control a pneumatic hook (which is often referred to as the “mezzoforte hook”) that prevents the expression pneumatic from fully opening or closing. The damper (sustain) pedal and una corda (soft) pedal are controlled by lock-and-cancel valves: holes in one track turn on the valve, and holes in an adjacent track turn it off, allowing the hole to have a continuing effect once triggered.

The “Crescendo forte” and “Crescendo piano” hole tracks produce slow crescendos (increasing dynamics). They are also controlled by lock-and-cancel valves. This means once a perforation of “Crescendo forte” is triggered, the dynamic will increase until it is cancelled by “Crescendo piano”. The “Forzando forte” and “Forzando piano” produce fast crescendos and decrescendos. They are controlled by a single continuous roll perforation, unlike lock-and-cancel valves, so the pneumatic is powered for the length of the perforation. When the crescendo and forzando tracks are not activated, a steady slow decrescendo results. In fact, a slow decrescendo is effectively constantly activated due to its connection to ambient air pressure.

To calculate the rate at which the crescendo or decrescendo operates, we examined original Welte T-100 test rolls, manuals and other sources [5,14]. A test roll contains a number of specific note and dynamic tests that allow a technician to adjust the fine regulation of the player piano. The slow crescendo, slow decrescendo, fast crescendo, and fast decrescendo rates are regulated by test 3 through 6 in the T-100 test-roll manual. Test 3 regulates the slow crescendo rate to the mezzoforte hook. Test 4 regulates the fast crescendo and decrescendo. Test 5 regulates the re-

---

4 The original technology used by reproducing roll companies to capture dynamic information from the live performer remains unclear. One theory suggests that the dynamics were translated into perforations from expression lines drawn by styli attached to pneumatics that were connected to expression regulators. Others have suggested a combination of electro-pneumatic valves and a dynamic rotor. A more detailed investigation of this question is beyond the scope of this paper but further information can be found in other sources [11,14].

Table 1: MIDI note number and corresponding perforation track information (encoded in the raw MIDI files).
<table>
<thead>
<tr>
<th>Crescendo type</th>
<th>Travel between</th>
<th>Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow Crescendo</td>
<td>min. to mezzoforte</td>
<td>1190</td>
</tr>
<tr>
<td>Slow Decrescendo</td>
<td>mezzoforte to min.</td>
<td>2380</td>
</tr>
<tr>
<td>Fast Crescendo</td>
<td>min. to mezzoforte</td>
<td>200 to 300</td>
</tr>
<tr>
<td>Fast Decrescendo</td>
<td>max. to min.</td>
<td>70 to 210</td>
</tr>
</tbody>
</table>

Table 2: Crescendo type and the time it takes to travel between two dynamic levels.

lease from fortissimo touch to piano touch. Test 6 regulates the combination of crescendo and decrescendo. These tests would have been used in conjunction with the judgement of a trained technician. Even as specified, they allow for a small range of results to the tests.

Based on this research, we chose rates for these parameters as shown in Table 2. We have chosen to model pneumatic pressure changes during execution of the slow crescendo and decrescendo as a non-linear function (one-pole filter), whereas that of the fast crescendo and decrescendo is modelled as a linear function. We apply the following constraints to map pneumatic pressures to MIDI velocity levels: the minimum velocity (softest level) is 35, mezzoforte (medium level) is 65, and the maximum (loudest level) is 90. In addition, the overall velocity of bass notes are assigned to be about 5 velocity levels lower than the treble notes.

5.3 Tempo Configuration

Most piano rolls formats are played with a speed written at the start of a roll by the manufacturer; however, Welte T-100 rolls do not have this indication. Test rolls and manuals do suggest that these rolls are to be played at single set speed, however Welte player instruments curiously come with a speed-adjustment lever. It has been theorized that this lever may have allowed for small adjustments needed by poorly regulated instruments. We deduce the general set tempo from an examination of test rolls and sources to be approximately 9.46 ft./min., but accept a potential range of ±0.5 ft./min.

Timings of notes in image-extracted MIDI files are expressed in delta-time ticks (pulses) that represent one pixel row in the original image. This allows notes in the MIDI file to be linked to their source locations in the scanned rolls. To translate roll speed into MIDI playback speed, we first set the initial MIDI tempo to 60 BPM. Then we calculate ticks (or pulses) per quarter note (TPQ, or PPQ, in the file header) by multiplying the roll speed with a factor that converts ft./min. into pixels/sec.: 

\[
\text{factor} = \frac{\text{feet}}{\text{min}} \times \frac{300 \text{pixel}}{\text{inch}} \times \frac{12 \text{inch}}{\text{feet}} \times \frac{1 \text{min}}{60 \text{second}} = 60
\]

So a speed of 9.46 ft./min. is used to set the TPQ value in the MIDI header to 568.

5.4 Paper Acceleration

The take-up spool on a player piano rotates at a constant speed, while the diameter of the spool increases as paper is wound around it. This causes an acceleration of the paper over time. Since the MIDI file time unit represents spatial distance on the roll, tempo changes are given in the MIDI file to emulate this acceleration. Based on our research and sources, we set this to be a 0.22% acceleration rate per foot [4]. The result is a more accurate tempo across the length of the roll.

5.5 Hole Extension

Another correction we apply concerns the effective length of paper holes over the tracker bar. The pneumatic valve is actually triggered before the paper hole completely lines up with the tracker bar hole, as enough of the ambient air pressure is generated by a portion of the hole. Thus the effective length of the hole is actually longer than it is on the paper, as illustrated in Figure 8. Compensating for this extension is especially important for the forzando piano and forte expression tracks because they are fast-acting and sensitive to small variations in time. We approximate the extra time the valve is open by measuring the length of a Welte-Mignon tracker bar hole and multiplying it by 0.75.

5.6 Audio Rendition

We provide high-quality audio transfers of the expressive MIDI files in SUPRA-RW by rendering them through Synthogy Ivory II Pianos in Steinberg Cubase. The resulting wav and m4a files should generate a result that compares favorably to a Welte T-100 roll played on an original Welte-Mignon pneumatic instrument.

6. CONCLUSION

In this paper, we describe the creation of the SUPRA-RW dataset which includes scanning, symbolic music extraction, pneumatic system modeling, and performance rendering via emulation. These procedures will apply with some variation (especially in emulation algorithms) to other reproducing roll formats and should be useful for those working to digitize large collections.

There is much work to do with reproducing (and non-reproducing) rolls. Our project will continue to study pneumatic expression mechanisms and we plan to create emulations for other roll formats. This will support online access to the thousands of rolls in Stanford’s collection and will grow the SUPRA database.
7. ACKNOWLEDGEMENT

We thank Anthony Robinson for contributing his roll scanner design and expertise to the project. Brett Swope of Swope Design Solutions built the scanner for this project. We thank Peter Phillips and Wayne Stahnke for their generous advice and consultation. The Stanford University Library staff provided expertise in many areas for which we would like to thank Tony Calavano, Doris Cheung, Frank Ferko, Nathan Coy, Jill Sison, Kristen St. John, Peter Crandall and Andree Clair Bonnepart. This project is funded partially through the generosity of an anonymous donor.

8. REFERENCES


