LBNL Report 51983
26-March-2003

Reconstruction of Mechanically Recorded Sound by Image Processing

Vitaliy Fadeyev and Carl Haber[†]
Physics Division
Lawrence Berkeley National Laboratory
1 Cyclotron Road
Berkeley, California 94720

Abstract

Audio information stored in the undulations of grooves in a medium such as a phonograph record may be reconstructed, with no or minimal contact, by measuring the groove shape using precision metrology methods and digital image processing. The effects of damage, wear, and contamination may be compensated, in many cases, through image processing and analysis methods. The speed and data handling capacity of available computing hardware make this approach practical. Various aspects of this approach are discussed. A feasibility test is reported which used a general purpose optical metrology system to study a 50 year old 78 r.p.m. phonograph record. Comparisons are presented with stylus playback of the record and with a digitally re-mastered version of the original magnetic recording. A more extensive implementation of this approach, with dedicated hardware and software, is considered.

1. Introduction

The preservation of mechanically recorded sound is an issue of considerable current interest [1,2,3,4]. Extensive recorded sound collections and archives exist world-wide. Many older mechanical recordings are damaged or are considered at risk for deterioration with time or due to further contact with a playback stylus. Some valuable recordings were only made as “instantaneous” transcriptions on cellulose acetate or cellulose nitrate and are particularly delicate. A method of extracting sound from these samples, which would do no further damage, is therefore attractive. Playback with a stylus only samples the portion of the groove wall in contact. Since better quality information may still reside in other parts of the groove cross section, a method of extracting information from any region of the groove is desirable as well. Furthermore, there is considerable interest in methods and strategies which would allow mass digitization of analog recordings for archival storage and distribution.

In the present work some techniques of digital image processing and precision optical metrology have been applied to the problem of extracting audio data from recorded grooves. In the methods discussed here, a large number of appropriately magnified sequential digital images of a groove pattern are acquired using an electronic camera or

---

[†] The authors can be contacted at CHHaber@LBL.GOV, and VAFadeyev@LBL.GOV

other imaging system. These images can then be processed to extract the audio data. Such an approach offers a way to provide non-contact reconstruction and may in principle sample any region of the groove. Furthermore, such scanning methods, if sufficiently fast and precise, may form the basis of a strategy for larger scale digitization of mechanical recordings which retains maximal information about the native media.

An example of one of a sequence of images is shown in Figure 1. It depicts a magnified groove field viewed with coaxially incident light and acquired with an electronic camera. The thin bright lines are the groove bottoms. This sample is taken from a 78 r.p.m. recording. The groove width at the record surface is about 160 microns. A 10 KHz tone would have a wavelength on the grooves between 40 and 100 microns at the minimum and maximum radii on the record respectively.

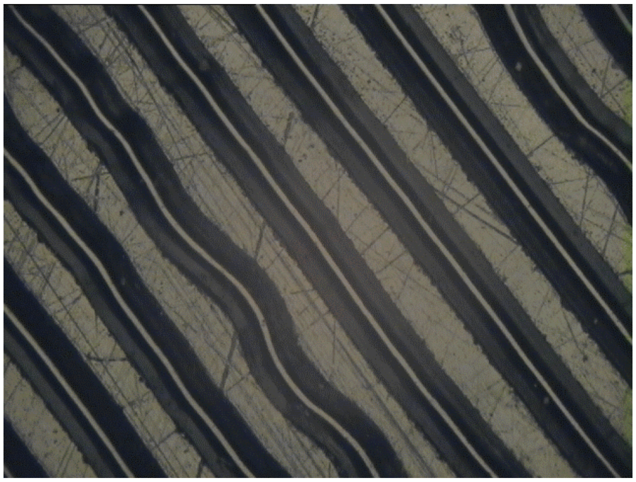


Figure 1: Micro-photograph of grooves on a 78 r.p.m. recording. Illumination is coaxial. Image size is approximately 700 x 540 microns.

Another example of an image is shown in Figure 2a. It depicts a three dimensional map of a groove field (including a large scratch) also from a 78 r.p.m recording, acquired with a confocal laser scanning probe [5]. This is a method which builds up an image frame by measuring a series of points across the surface. The data were acquired on a 4 micron square grid. The vertical resolution is 100 nanometers.

A third example of an image is shown in Figure 2b. It depicts data also acquired with a confocal scanning probe from the surface of an Edison Blue Amberol cylinder. The data were acquired on a 4 micron square grid, again with 100 nanometers vertical resolution.

With available metrologic methods sufficient resolution is achievable to measure the full range of groove movements in mechanical recordings. Typical required resolution is in the sub-micron range. These methods are also sensitive enough to measure some of the effects of wear and damage which may then be corrected for in image analysis and processing. Recorded disks with lateral groove movement may be reconstructed in either two (2D) or three dimensions (3D) as appropriate for the condition of the sample. Cylinders and disks with vertical “hill and dale” undulations would be scanned by 3D methods such as the confocal process shown in Figures 2a and 2b or stylus profilometry

[6]. In the latter case, the surface is contacted, but by a low mass stylus incapable of doing any damage to the sample. With 2D cameras, lighting options can affect the imaging strategy. Using coaxial illumination it is most natural to image the groove bottom or the top edge. With other types of illumination alternative aspects of the groove could be imaged.

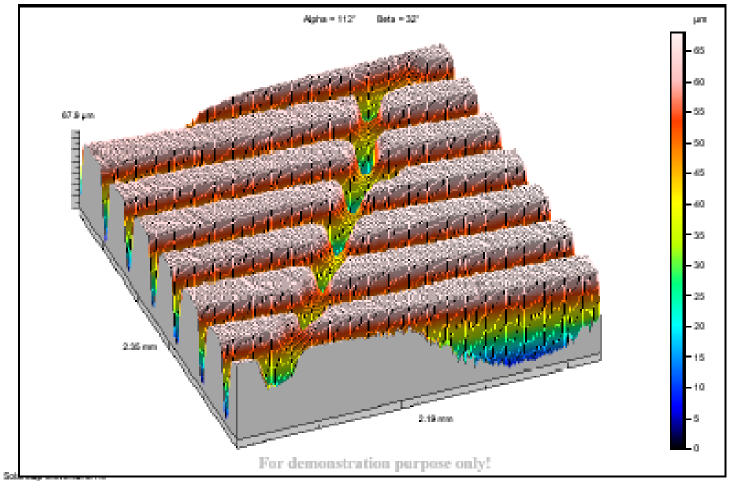


Figure 2a: Image data of scratched groove field on a 78 r.p.m. recording acquired with a commercially available laser confocal scanning probe [5]. Size is 2.35 x 2.19 millimeters.

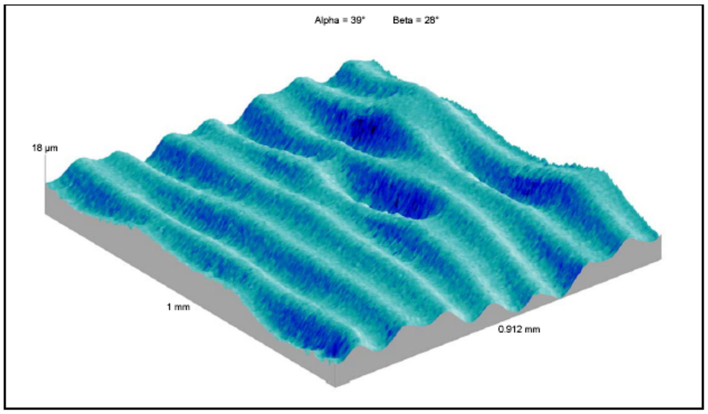


Figure 2b: Image of a region of the surface of an Edison Blue Amberol cylinder. The field size is 1 x 0.91 mm and is also acquired with a commericially available laser confocal scanning probe [5].

With available computing tools, the image data generated in the scan of a phonograph record or cylinder surface can be handled efficiently. Image analysis methods can be applied to model the local groove shape by appropriate mathematical functions or discrete series. Using these models it is then possible to calculate the motions a stylus would have made when it passed along these grooves and filter or remove the effects of scratches, dirt, and wear through image analysis and transformations.

Reconstruction of mechanically recorded sound by image processing should be distinguished from the use of laser or light beam based turntables [7,8]. The latter either replace the stylus with a reflected light beam or scatter light off a large mechanical stylus in contact with the groove. Both are therefore susceptible to the effects of dirt, damage, and wear. The methods discussed in this paper always rely on the analysis of digitally imaged frames of the recording.

Reconstruction of recorded sound by image processing can be applied to broken or warped media and is not particularly sensitive to material composition or color.

The methodology applied in this work is derived, in part, from long standing analysis methods used in high energy particle and nuclear physics to follow the trajectories of charged particles in magnetic fields using position sensitive radiation detectors [9,10,11]. It also benefits from the development of tools used in the characterization of semiconductor devices and in automatic visual inspection.

In the body of this paper, the reconstruction of data from lateral groove recordings using 2D optical metrology is emphasized. However, the framework and most of the issues also apply to 3D and scanning methods as well. A proof of concept, applied to a 78 r.p.m. disk manufactured around 1950, was carried out and is described. Some considerations leading to a practical implementation of this method are addressed.

This paper is organized in the following way. The concepts and principles of the method are discussed in Section 2. Section 3 presents a description of the proof of concept test and a discussion of the results of that test. Section 4 describes features of a dedicated system for reading and reconstruction of recorded media. Final conclusions are summarized in Section 5.

2. The Imaging Method

In this discussion the focus will be on applications to measurements of lateral groove recordings. In order to establish a basis for digital image processing and precision metrology it is useful to summarize the relevant mechanical properties of lateral recordings. The specifications for modern disk records are in print [12]. The corresponding specifications for the now obsolete, 78 r.p.m. technology are out of print but can be found in various sources as well [13,14,15].

The playback stylus signal is proportional to its transverse velocity which is due to the lateral groove movement. Signals are compared on the basis of amplitude rather than power.

$$dB = 20\log\left(\frac{v}{v_{REF}}\right) \qquad (1)$$

where v is the stylus velocity and v_{REF} is some defined reference level, discussed below.

For a sinusoidal modulation, the signal amplitude read with a stylus will be maximal at the zero crossings of the groove. The maximum lateral displacement of the groove corresponding to a zero crossing velocity v_{MAX} is,

$$A_{MAX} = \frac{v_{MAX}}{2\pi f} \qquad (2)$$

where f is the frequency of the recorded tone. From an image of the groove pattern, the lateral displacement of the groove with respect to the un-modulated trajectory is measured on a sequence of points. The measurement of stylus velocity, at each point, is extracted from this displacement waveform by numerical differentiation.

Equation 2 states that for constant stylus velocity the maximum groove displacement depends inversely upon frequency. In most recordings the lower frequency sound levels are deliberately attenuated to increase the range of signals which will fit in the groove spacing allocated. In addition, higher frequency sound levels are often boosted to overcome the drop off of equation 2 and to raise the signals above a high frequency noise floor. This “surface noise” is inherent to the mechanical recording process. For recordings before the early 1950’s this process of equalization was not standardized, but many of the curves have been tabulated [16]. (A proper interpretation of a recording reconstructed either by mechanical playback or optically should have the equalization compensated for.)

Before defining the basic physical parameters of the mechanical recording it is useful to summarize definitions for the various noise sources which may be encountered. In the literature of mechanically recorded sound it is common to refer to effects which are either random, or systematic and distorting, as “noise”.

1) <u>Surface noise or hiss</u>: This is a random high frequency noise which is due to some imperfections in the groove side surface. These imperfections are inherent in the material used to make the record, but may increase due to effects of age and wear. In general a stylus will be very sensitive to the groove surface. Depending upon the imaging strategy applied (such as illumination), an optical approach may be less sensitive to the surface quality. The frequency spectrum of surface noise remains flat up to the highest audible frequencies. Since the surface noise is random its measure would be the standard deviation of the surface noise velocity or amplitude distribution about an un-modulated groove trajectory.

2) <u>Transient impulse noise or “clicks and pops”</u>: This type of noise is due to discrete imperfections or damage sites along a groove, such as a scratch. The noise pulses are of short duration but may have large amplitude. They occur at random times but are typically isolated. From an imaging approach such defects are resolved and can be handled by basic image processing methods. If they are removed from the data then the lost portions can be estimated from the surrounding groove profile. This interpolation is analogous to what is done in some noise reduction systems which work on the mechanically played back data. In an imaging approach more information is available about the offending structures. They can be fully visualized

and this may aid in their removal. Any dynamical effect on the stylus motion which persists after the impulse [17], including a complete skip, will also be absent in using imaging methods.

3) <u>Wow and flutter</u>: These are low frequency effects which are probably due to variations in motor speed, off axis position of the center hole in the record, acoustic feedback, and non-circular groove shape to indicate a few. The quoted frequency ranges are below 6 Hz for wow and between 6 and 30 Hz for flutter. These are not actually noise in a technical sense but rather systematic distortions which affect the performance of the system. They are typically characterized as some maximum allowed deviation from an un-modulated groove rather than by a statistical measure. In an imaging approach these effects are essentially irrelevant since they are either absent or can be removed through shape parameters in the analysis.

The main mechanical parameters of interest for digital imaging and precision metrology are listed below. Some relate to Figure 3 which depicts a groove profile. When relevant they are defined at a specific frequency (1000 Hz) where equalization effects are generally not an issue.

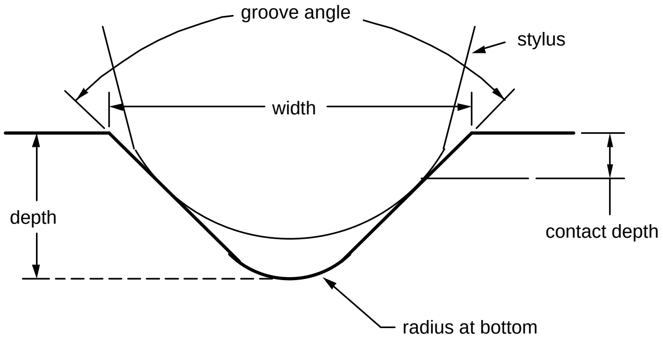
1) <u>Groove width</u>: distance across the top of the groove (defined in Figure 3).
2) <u>Groove spacing</u>: center to center distance between two adjacent grooves.
3) <u>Grooves per inch</u> (G_d): the number of grooves cut in the surface per radial inch.
4) <u>Reference signal level</u>: the peak transverse velocity used to set a baseline for the recorded signal. This quantity is in principle arbitrary but is key to defining the noise and dynamic range discussed in the literature.
5) <u>Maximum groove amplitude</u>: the maximum displacement of the groove from an un-modulated path.
6) <u>Noise level below reference level (signal to noise ratio)</u>: noise levels or limits are usually expressed as dB below the reference signal. This is taken to mean the standard deviation of any random noise source, such as the underlying surface noise source discussed above, or the maximum allowed deviations due to the low frequency systematic effects.
7) <u>Dynamic range</u>: a measure of the range of audible signals up to the maximum peak recorded signal level, defined here with respect to the noise level at 1000 Hz
8) <u>Groove amplitude at noise level</u>: the maximum amplitude deviation from a signal free path corresponding to the noise level in item 6) above and defined in equation 2.
9) <u>Maximum and minimum radii</u>: the respective radii at which audio data is specified to begin (R_{MAX}) and end (R_{MIN}).
10) <u>Area</u>: the area covered by audio data.

$$Area = p\,(R_{MAX}^2 - R_{MIN}^2) \qquad (3)$$

11) <u>Total length</u>: the path length along a complete groove between the two radial extremes.

$$L = G_d \times Area \qquad (4)$$

These parameters are presented in Table 1 for the 78 r.p.m. coarse and 33 1/3 r.p.m. ultrafine groove technologies. The units used follow the past conventions where applicable.



parameter	78 rpm coarse	33 1/3 ultrafine
width	0.006 - 0.008	0.001
depth	~0.0029	~0.0006
contact depth	0.0008	0.0004
radius	0.0015-0.0023	0.00015
angle	82 - 98	87 - 92

units are inches or degrees

Figure 3: Section of stylus in groove for the 78 r.p.m. coarse groove and 33 1/3 r.p.m. ultra-fine groove technologies.

Parameter	**78 r.p.m., 10 inch**	**33 1/3 r.p.m., 12 inch**
Groove width at top	0.006-0.008 inches	0.001-0.003 inches
Grooves/inch G_d	96-136	200-300
Groove spacing	0.007-0.01 inches	0.0033-.005 inches
Reference level peak velocity@1KHz	7 cm/sec	7 cm/sec (0.0011 cm)
Maximum groove amplitude	0.004-0.005 inches	0.0015-0.002 inches
Noise level below reference, S/N	17-37 dB	50 dB
Dynamic range	30-50 dB	56 dB
Groove max amplitude at noise level	1.6 - 0.16 µm	0.035 µm
Maximum/Minimum radii	4.75/1.875 inches	5.75/2.375 inches
Area containing audio data	38600 mm^2	55650 mm^2
Total length of groove	152 meters	437 meters

Table 1: Basic mechanical parameters of some lateral groove recordings.

From the mechanical parameters described above, the basic requirements of a measuring and data processing system can be derived. The fundamental requirement is on measurement resolution and accuracy. In the context of electronic imaging, resolution refers to the statistical error on the measurement of a point and is due to the pixel size plus any additional effects of the optical chain. The pixel size effect implies the resolution will be

$$s = \frac{\text{width}}{\sqrt{12}} \qquad (5)$$

where “width” is the image size projected onto one pixel. Accuracy refers to the possibility of discrete shifts of data between pixels and is attributed to timing or phase jitter in the electronics chain which reads out the image. Requirements on both are essentially set by the intrinsic record noise level. The metrology process stands to add additional noise to the sample if the accuracy and resolution are worse than the intrinsic noise. The effect is actually magnified since audio extraction is done by differentiating the measured groove pattern. This magnification is dependant upon the process used to differentiate and the sampling rate. In addition, poor resolution will smear the waveforms leading to a loss of information.

Having determined the basic measurement accuracy and resolution an appropriate imaging system can be designed. There will be a set of tradeoffs between magnification, field of view, number of pixels, and data rate. Such optimizations are the subject of various published discussions [18]. In surveying the surface of a record, the imaging system will scan over a number of positions. The mechanical motion may induce an additional shift error. If adjacent frames overlap, the data can be used to correct for this at the expense of a larger data set. Similarly, for fixed sampling rate, higher magnification can lead to improved resolution at the expense of this redundancy. A specific example of a reasonable camera format and optical chain follows in the discussion of test results in this paper.

For a scan with an electronic camera, typical data size captured in a single monochromatic frame is ~1 Megabyte. A reasonable field of view (see next section for details) for the measurement of the groove pattern is the 700 x 540 microns (0.378 mm^2) shown in Figure 1. From Table 1, the recorded surface area on a 10 inch, 78 r.p.m. coarse groove sample requires about 10^5 such fields for complete coverage, assuming the data is efficiently used. About 10^6 fields would be required if just a single groove segment is scanned at any one step. The total data set generated is then 100-1000 Gbytes before processing. While large this is still manageable. In an audio restoration application there is no strong requirement on the reading speed of the measurement system except overall throughput, if many disks are to be surveyed. As an example to consider, a system which scans the record surface in real-time with an electronic camera would generate between 0.5-5 Gbytes/second. Again, such a data stream might be handled if sufficiently parallelized. It is common in machine vision applications to follow the sensor directly with a fast digital signal processing module which can usually reduce the data size considerably before transfer to the host computer.

Imaging the grooves includes important advantages over standard stylus playback methods. These are stated or restated here.

1) Some old recordings, which are of historical, scholarly, or perhaps commercial value, are considered too delicate or otherwise compromised to read with a stylus. The imaging methods are essentially or totally non-contact.

2) Old recordings are often damaged by dirt and scratches or worn down. Using image

analysis methods and lighting options including intensity, angle of incidence, and wavelength, the damage and debris may be removed (filtered) since they may be recognizable as different from the groove structure. The effects of wear may be overcome by sampling regions of the groove away from where a stylus would run or using interpolation or correction methods. Included is the use of the groove bottom which may be free from the effects of scratches and wear. To an extent this is similar to the use of a truncated or specially shaped mechanical stylus but more general. An imaging method may be less sensitive to material characteristics which contribute to surface noise in stylus play back.

3) Existing methods for improving the quality of recorded sound utilize filters and reconstruction methods which are applied to the already played back signals in the time or frequency domain. The imaging method performs a reconstruction and filtering in a <u>spatial or image domain</u>, where the noise actually originates, through the acquired image alone.

4) Dynamic effects of damage or debris which may excite a resonant or disruptive response in a mechanical stylus and cartridge [17] are absent in an optical or precision profile-metric reading of the grooves.

5) These methods may be applied to any medium upon which the recording is made including vinyl, shellac, wax, metals, and other opaque or transparent plastics. Broken or warped media may also be reconstructed.

6) Various sources of distortion and noise such as wow and flutter, tracing error, pinch effects, and tracking errors, are either absent in this approach, or can be resolved as simple geometrical corrections in the analysis.

7) Systems which read records or cylinders by means of just a reflected light spot must include a mechanical mechanism for following the groove. If image data are acquired the data itself is aligned in software simply by matching adjacent frames and through use of encoder read back and stage indexing.

8) Some recordings may only exist in the form of metal stampers. While there are methods to play back some of these, using a special stylus, play back through imaging methods is no more difficult than groove imaging. An example of a 2D image taken from a stamper is shown in Figure 4. The focus is set to the peak height which corresponds to the groove bottom.

After an image based reconstruction any of the existing digital noise reduction or re-mastering techniques can also be applied to further enhance or otherwise alter the sample.

Imaging the grooves also introduces an additional set of technical issues which should be considered to optimize the performance of the method. These are stated or restated here.

1. The resolution and accuracy of the entire imaging and data chain must be sufficient to

measure the undulations without introducing excess noise or smearing of the signals. Field of view and magnification tradeoffs must be considered.

2. The mechanical indexing must be accurate enough to allow simple correlation of adjacent frames.

3. An optimized digital signal processing algorithm must be used to determine the modeled stylus velocity through the process of numerical differentiation without adding excess noise.

4. Sufficient points (sampling rate) must be found along the groove for correct modeling.

5. Processing time and data storage requirements increase if less pre-processing is done in the initial acquisition chain and/or more points are sampled or higher magnification is used. The methods discussed here are not necessarily meant to play the recording back in real time. Rather they extract maximal information from the sample to enable substantial processing if required. With improvements in technology the potential for faster read back and processing grows.

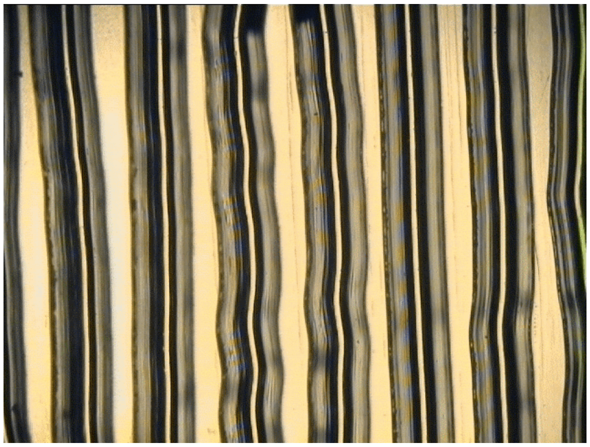


Figure 4: Image of a metal stamper.

3. A Test of the Imaging Method

In this section a test of an optical reconstruction method is described. The test was based upon existing or easily available tools and methods. Little of the optimization discussed above was carried through. For this reason, the results should not be considered as a definitive indication of the power of these methods but rather as a starting point for further refinement or development. In light of these contingencies, the results obtained were judged quite satisfactory as compared to a stylus playback of the same source material.

To test the procedure a general purpose optical metrology system with digital image processing capabilities was used. The device was the “Avant 400 Zip Smart Scope” which is manufactured by Optical Gauging Products [19]. A view is shown in Figure 5.

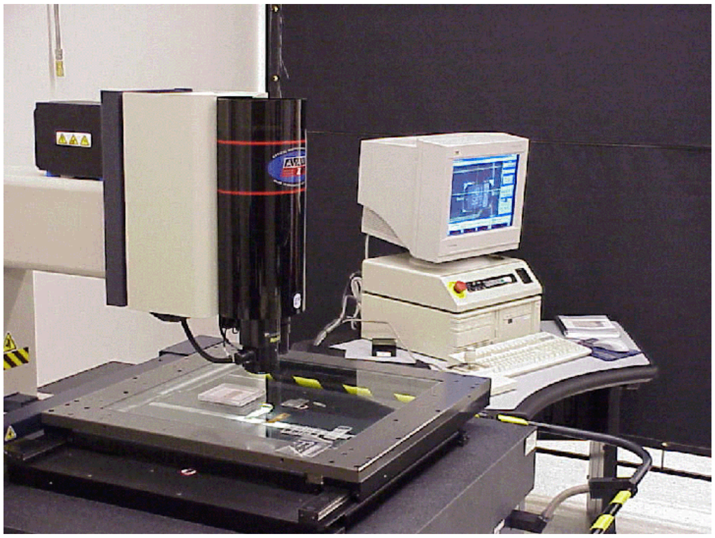


Figure 5: Optical Gauging Products Avant 400 Zip Smart Scope metrology system.

It consists of a video zoom microscope and a precision X-Y table. The accuracy of motion in the X-Y (horizontal) plane over the distance L (mm) is (2.5 + L/125) microns. The video camera had a CCD 6.4 mm x 4.8 mm containing 768x494 pixels of dimension 8.4 x 9.8 microns. With appropriate lenses installed it imaged a field of view ranging between approximately 260 x 200 microns and 1400 x 1080 microns. The optical system could travel in Z (vertically) with an accuracy of (4.0 + L/150) microns.

The system included a software package [20] containing a selection of image recognition tools useful for metrology. It could also be programmed to seek out known features on a part and repeat a measurement series at the user's discretion. Heights could be measured with a focusing criterion.

A typical 78 r.p.m. groove has a cross-section with a flattened bottom (Figure 3). The tip of the stylus glides along the sides of the groove. This test used illumination coaxial with the optics to accent the bottom part as possibly less worn. Magnification was chosen to give a 700 x 540 micron FOV. Each pixel of the CCD then corresponded to 0.91 x 1.09 microns on the record surface. Based upon Equation 5, the resolution of this system was 0.26 x 0.29 microns.

To carry out this test a particular procedure was implemented which was convenient to the available tools. A program was written which determined the disk center and followed the groove trajectory spiraling inwards. A sequence of images like Figure 1 were acquired and processed. In each image frame the two sides of the groove bottom were measured with an edge finding algorithm. In this case the edges were the light to dark transitions between the illuminated groove bottom and the sloping walls. Point pairs were taken every 8 microns along the groove path. This corresponded to a sampling frequency of 61.3 kHz at a groove radius of 60 mm. Raw data, consisting of coordinate pairs for these points, were written to a file. Dust or debris was generally skipped since no edge would be found near the groove path. In this way some limited image processing, and therefore noise reduction, occurred already. The raw data was then processed offline. This processing consisted of the following steps.

1. Data from each frame was merged into a global polar (R,ϕ) coordinate system. Here R is the distance from the disk center and ϕ measured the azimuthal position from the starting point.

2. Using the position correlations of data points on each side of the groove bottom, certain spurious points were removed from the data (filtering).

3. The good point pairs on either side of the edge were averaged. A sample of the data before and after this step is shown in Figures 6a and 6b.

4. Small mechanical shifts, seen in Figure 6, existed due to movements of the mechanical stages. Typical shifts were ~1 micron. Data were acquired in overlapping segments and this redundancy was used to remove the shifts.

5. The variation of the R coordinate with ϕ for the measured points is the result of the audio signal, the natural groove spiral shape, the effect of the disk center determination uncertainty, and (possibly) the manufacturing distortions, such as "wow". To minimize all sources, except the signal, the data were fit to the functional form $R' = R_0 + C*\phi' + A*\sin^2(\phi' + \phi_0)$. Here the R' and ϕ'are different from R and ϕ in that the former pair are recalculated from the latter by taking into account the proper disk center. The difference between R' for an individual point and the function shape gave the "signal" due to the recorded sound. The "wow" effect was attenuated by the $\sin^2$ function.

6. The remaining audio waveform was differentiated numerically to obtain the corresponding stylus velocity. This was done by fitting groups of 15 points, centered about each chosen point, to a 4th order polynomial and taking the analytic derivative of the fitted function. This method may not be optimal for audio data with some continuous noise present but certainly works well enough for the purpose of this test. It was seen to work better than a simple two point difference derivative and slightly better than fits to fewer points or to lower order polynomials.

7. The data were re-sampled at the CD standard frequency of 44.1 kHz and converted into WAV [21] format.

The speed of the scanning procedure was determined by the specifications of the metrology system used. In this case it took around 50 minutes to scan 1 second of recorded data. The system used was general and not optimized for this task. A dedicated system (discussed in Section 4) could be dramatically faster. It would be incorrect to conclude that the rate in this test represents the real achievable scanning rate.

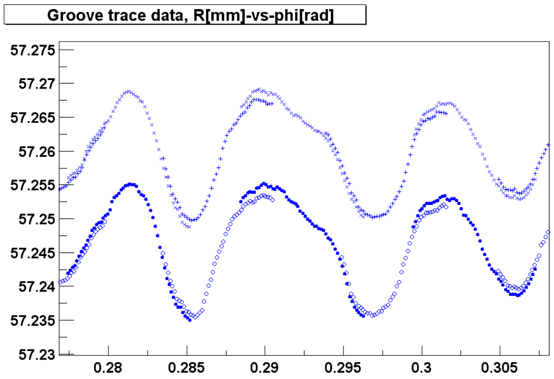


Figure 6a: A sequence of data points taken along the two imaged edges of the groove bottom plotted as a radial coordinate in mm (vertical axis) versus ϕ, an angular coordinate in radians (horizontal axis). A series of overlapping camera frames are shown. The discontinuity (radius shift) between adjacent overlapping intervals is the issue addressed in step (4) of the data processing.

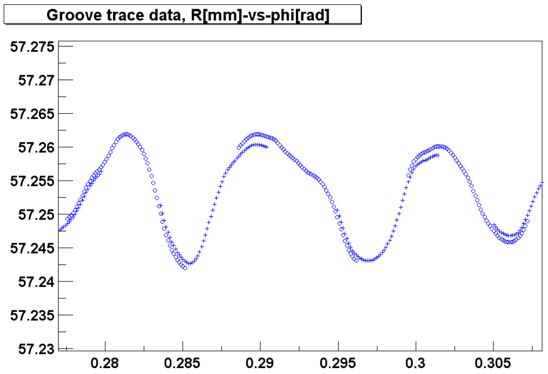


Figure 6b: Each point pair of Figure 6a is averaged across the groove bottom to form a single data stream. This redundancy also results in some smoothing of the data.

The test was performed on a 78 r.p.m. phonograph record which was manufactured around 1950 [22]. This recording was chosen at random and was not in particularly good condition. Not attempt was made to clean it. Figure 1 is an actual field from the surface of the record used in this test. The result of the scans and reconstruction was a 19.1 second sound clip. A 78 r.p.m turntable [23] and stylus [24] was also used to play the same record back through a commercial phono pre-amplifier [25] and PC sound card [26]. Again the record was not deliberately cleaned. Both the optically reconstructed clip and the mechanically played clip were displayed with commercial audio editing software [27]. The phono pre-amplifier used applied the standard RIAA equalization curve to the audio signals. To compare on an equal basis, the actual frequency response of the pre-amp was measured and applied to the optical data using a software equalization tool in the audio editing package [28]. No additional noise reduction process was applied to the digital audio files. Since this recording was made around 1950, the studio master was recorded to magnetic tape. In this particular case the tape masters still exist and this recording was recently re-issued on compact disk [29], re-mastered off the original tapes [30]. Because of this it is also possible to compare the optical and mechanical playbacks of the record to a high quality studio version. In the case of the compact disk version, the equalization used was not determined. All the clips can be accessed from the Web [31].

A set of figures is presented here to compare the different versions. These were displayed with the commercial audio editing software. Figure 7a, b, and c show the full 19.1 seconds for the optical, the mechanical, and the studio versions respectively. Figures 8a, b, c show a portion of the sound of duration 40 ms beginning at 3.261 seconds from the start of the optical clip for the optical, the mechanical, and the studio versions respectively. This segment contains rich audio data. Figure 8d is the audio data as read directly from the record before numerical differentiation into a velocity waveform. Figures 9a, b, and c are of the same description as Figure 8 except they are taken from a musically quiet portion of the recording (beginning at 17.661 seconds). Figures 10a, b, and c show an averaged Fast Fourier Transform (FFT) spectrum for each of the optical, the mechanical, and the studio versions respectively,

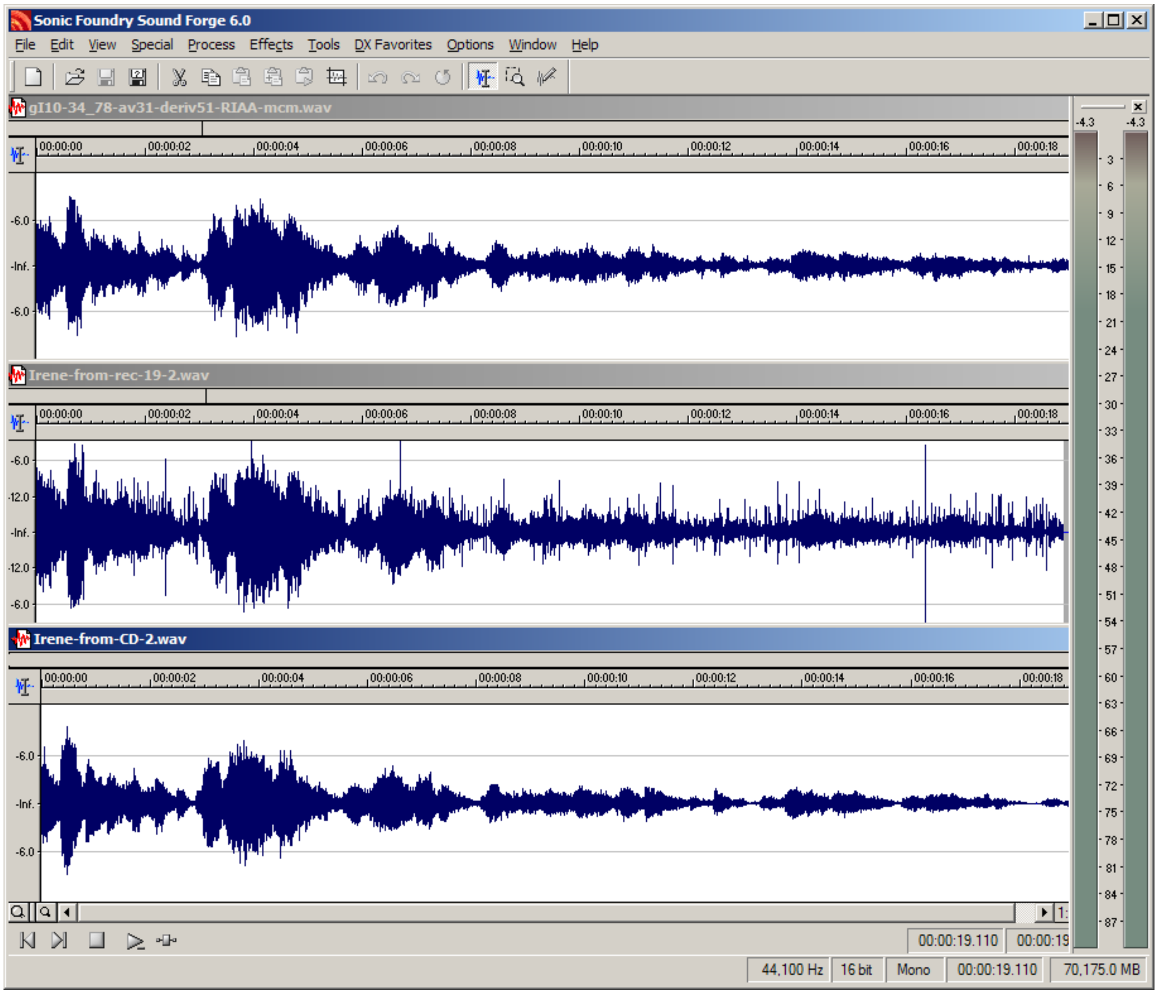


Figure 7: Full sound clip of 19.1 seconds from a) optical reconstruction (top trace), from b) stylus playback (middle trace), and from c) new CD release studio version (bottom trace).

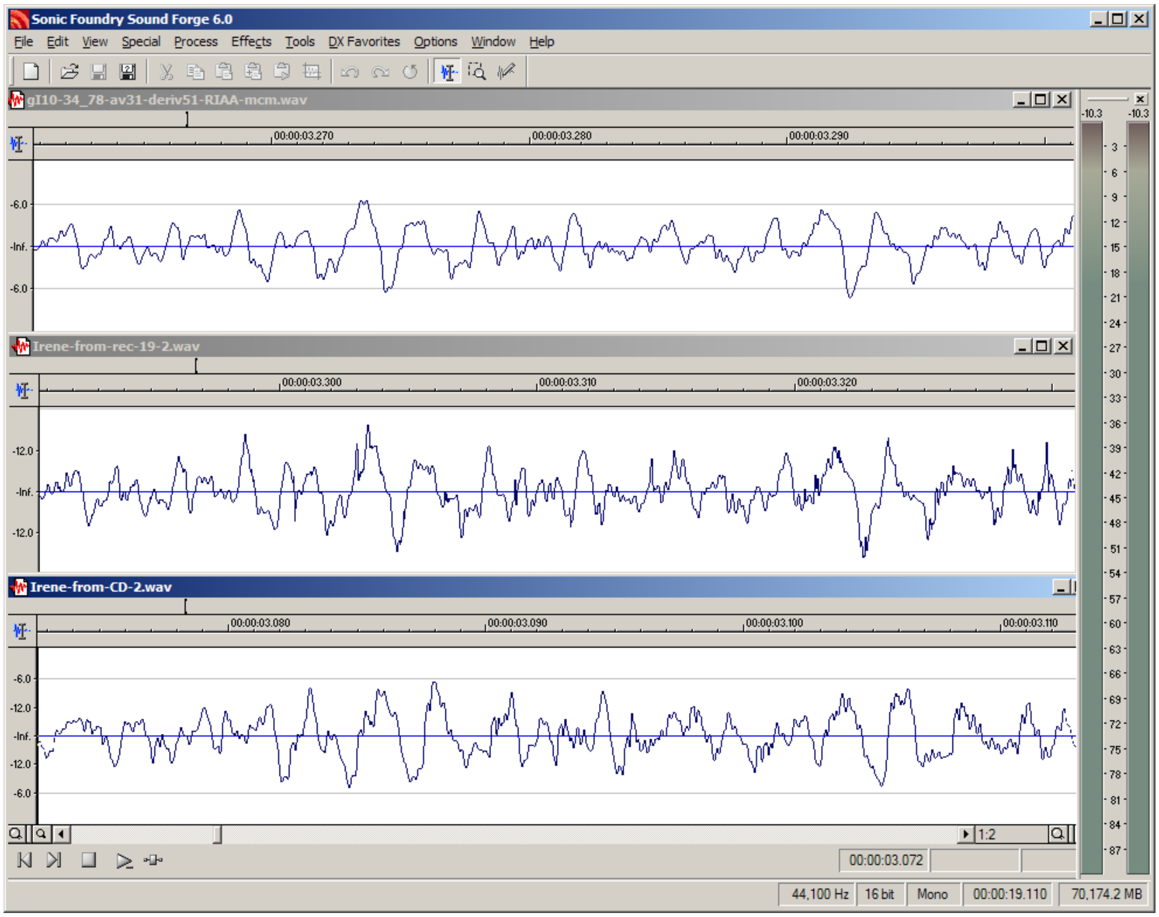


Figure 8: An expanded version of Figure 5 beginning at 3.261 seconds from the start of the optical sound clip and lasting 40 ms. The top a) is optical, middle b) from stylus and bottom c) from CD.

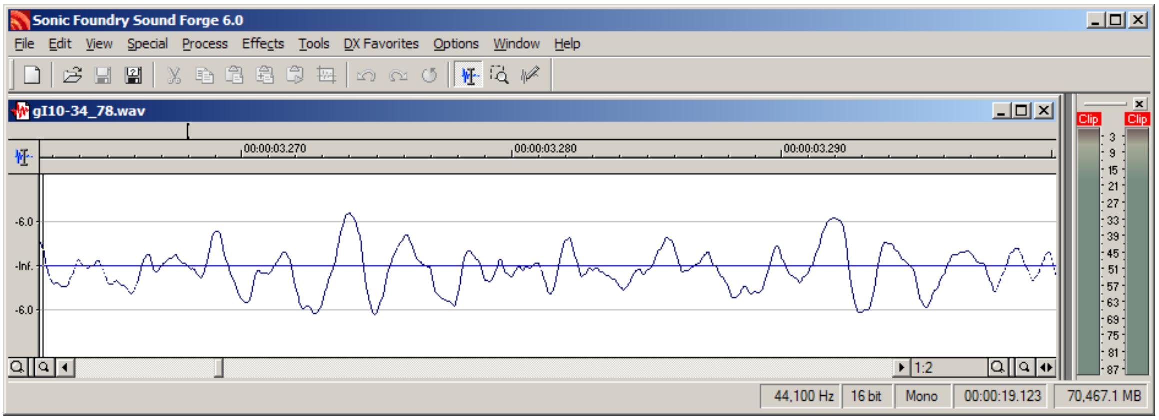


Figure 8d: The actual waveform of the groove amplitude versus time before numerical differentiation. The segment matches that of Figure 6a.

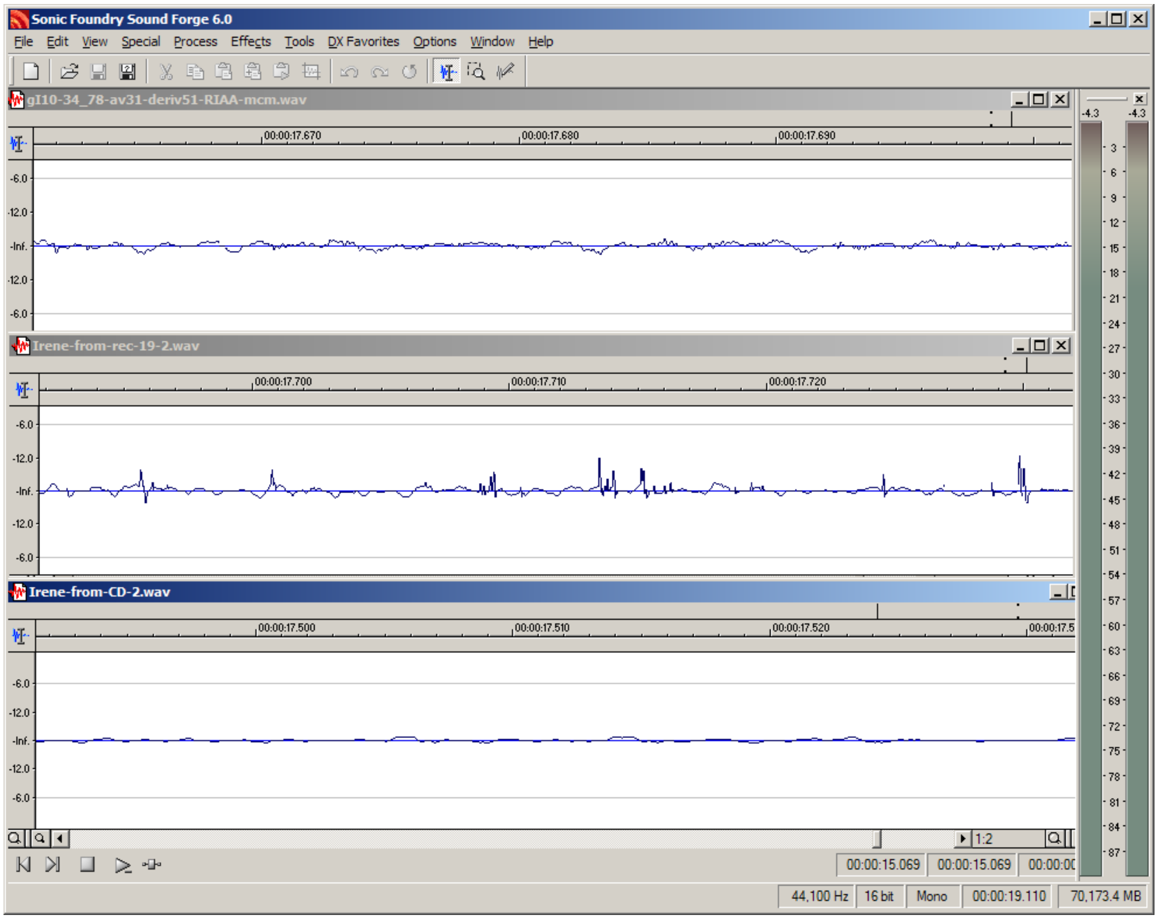


Figure 9: A relatively quite portion of the sound clip, beginning at 17.661 seconds from the start of the optical data. The absence of the sharp noise transients is clear in a) (top) from the optical read and c) (bottom) the CD version, but is seen in b) (middle) the stylus version.

From these results a number of conclusions can be drawn.

1. The fine waveform structures of all samples are qualitatively similar. This is particularly true of the optical and mechanical samples.

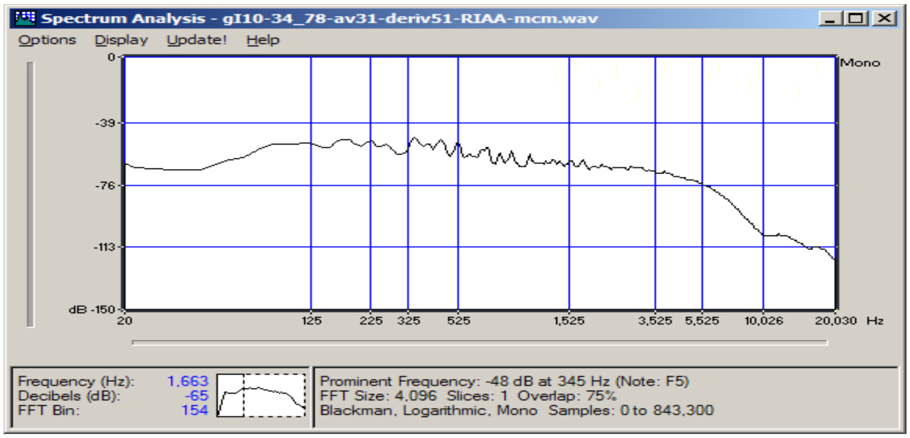
2. The optically read sample contains far fewer sharp noise features (clips or pops) than the mechanically played sample and also contains higher amplitudes (with respect to the broadband noise) in the musically rich segments.

3. The continuous noise level heard in the sound clips is lower in the optical sample than the mechanical sample.

4. A background continuous noise (hiss) is present in the optical sample. The hiss is also slightly modulated by a signal at about 4 Hz. The origin of this is not completely known but it may be related to the particular differentiation algorithm, imaging fluctuations in the edge finding process, or to a latent physical feature of the record

itself. A hiss signal is also present in the groove shape data before differentiation which may underlie the signal heard in the differentiated audio clip.

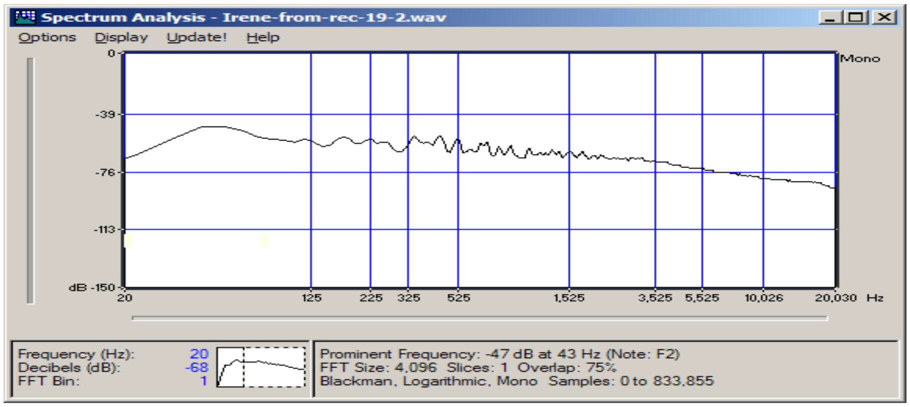
5. The studio version is, as expected, of higher quality than either the optical or mechanical sample. However the optical and mechanical sample were not subjected to any further digital noise reductions or re-mastering procedures.



Frequency (Hz): 1,663
Decibels (dB): -65
FFT Bin: 154

Prominent Frequency: -48 dB at 345 Hz (Note: F5)
FFT Size: 4,096 Slices: 1 Overlap: 75%
Blackman, Logarithmic, Mono Samples: 0 to 843,300

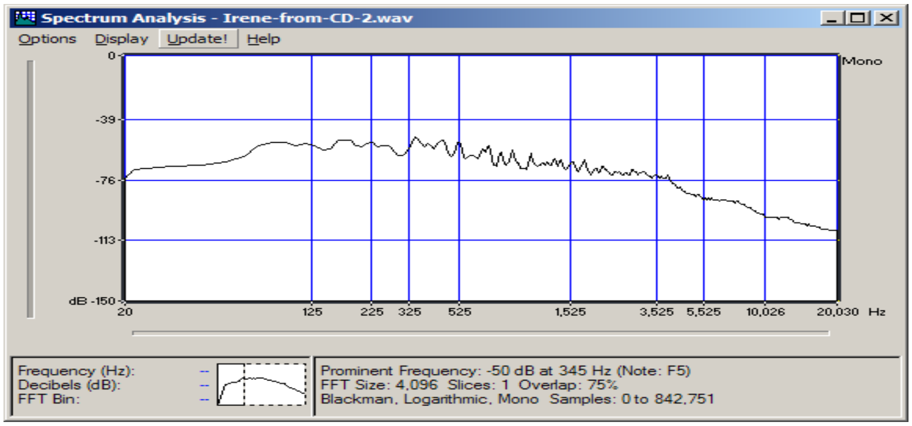
Figure 10a: FFT spectrum of the optical sound clip.



Frequency (Hz): 20
Decibels (dB): -68
FFT Bin: 1

Prominent Frequency: -47 dB at 43 Hz (Note: F2)
FFT Size: 4,096 Slices: 1 Overlap: 75%
Blackman, Logarithmic, Mono Samples: 0 to 833,855

Figure 10b: FFT spectrum of the stylus played back sound clip.



Frequency (Hz): --
Decibels (dB): --
FFT Bin: --

Prominent Frequency: -50 dB at 345 Hz (Note: F5)
FFT Size: 4,096 Slices: 1 Overlap: 75%
Blackman, Logarithmic, Mono Samples: 0 to 842,751

Figure 10c: FFT spectrum of the CD version sound clip.

A number of directions remain open for further development using these particular methods and tools. These include increased sampling rate, variations on the edge finding strategies, and alternate digital signal processing algorithms. For example, if every pixel had been used as a sampling point, these images would have yielded a 400-500 KHz sampling rate.

4. Design of a Specialized Machine

In a refined version of this technique an appropriate imaging or scan technology would be selected based upon required precision and rate. For a 2D approach, an optimized optical chain would be applied including appropriate magnification and camera performance to minimize resolution and accuracy effects. A dedicated X-Y-Z-Θ (rotating) movement system would be used to enable faster access to the medium. Other options could be applied as well such as light color, angle, and intensity as discussed above. In a 3D approach, special mechanical scanning methods could be developed to read cylinders rather than disks by coupling to a rotating stage assembly and using appropriate scanning and imaging methods. Custom designed image analysis would be applied to extract the maximum information from the groove data. A set of dedicated image based noise filtering and groove surface reconstruction routines would be developed. For broken or damaged media special reconstruction software would also be used to reassemble the tracks. The final tool could be a sound preservation workstation comprising hardware and software designed to handle the various cases presented. This workstation could include a suite of metrology tools to be applied as appropriate.

With dedicated processing hardware, optimal use of the imaged field, and faster large format cameras and/or multiple cameras or scanners, a dramatic increase in the speed of this process could be achieved. The raw image data produced by the scan of a 78 r.p.m. record with an electronic camera is in the range of 100-1000 Gbytes if no pre-processing is done. If the record was scanned in real-time this corresponds to 0.5-5 Gbytes/second. In a 3D profile scanning approach, a complete map of the surface of a cylinder on (for example) a 4 micron square grid contains 760 million points. Depending upon the storage format adopted this could represent up to a few Gbytes of data per scan. While the corresponding data rates for real time scanning are large, they are not unusual in the context of modern high speed on-line data processing as used in the triggering and data acquisition systems of high energy physics experiments [33,34]. There-in considerable parallelism is often applied to the incoming data streams and preliminary data reduction is done in dedicated hardware or firmware.

5. Conclusions

The basis for digital image processing and precision metrology, applied to mechanical sound reconstruction, has been described. A non-contact reconstruction of analog audio data from groove recordings has been demonstrated in a simple proof of concept test. The quality is already better than that achieved by stylus playback with good components before any other digital noise reduction methods are applied. Considerable options exist for improved image processing and further noise filtering. Application specific hardware and software could lead to significant reductions in the time required to perform a scan and further improvements in data quality. These attributes may lead to a real advantage in the preservation of endangered audio media of historical or other value.

6. Acknowledgements

This work was supported by the Laboratory Technology Research Program (SC-32), within the Office of Science, U.S. Department of Energy under Contract No.DE-AC03-76SF00098.

The authors wish to thank C.Kniel, H.Sartorio, S.Rosen, and M.Granados, of the Lawrence Berkeley National Laboratory Technology Transfer, Legal, and Procurement Departments for support and assistance in this work. The authors wish to thank R.Waxler and H.Spieler for useful comments, suggestions, and references. The authors wish to thank S.Teicher for providing record samples. The authors wish to thank Solarius Development Inc, for providing confocal scan images. The authors wish to thank R. Koprowski for providing samples and useful comments.

7. Note

Following the completion of the work described here-in the authors became aware of two related references. A paper presented at the 20th International AES Conference in Budapest in 2001 [34] describes an approach to archiving lateral recordings based upon image capture on photographic film followed by 2D scanning of the film. A web site [35] describes an attempt to read a stereo disk using a desk top scanner.

8. Disclaimer

See this link for a disclaimer.
http://www-library.lbl.gov/teid/tmRco/howto/RcoBerkeleyLabDisclaimer.htm

9. References

[1] Hart, M., **J.Audio Eng. Soc**, Vol. 49, No. 7/8,2001 July/Aug
http://www.aes.org/journal/suppmat/hart_2001_7.pdf

[2] http://www.saveoursounds.org

[3] http://lcweb.loc.gov/folklife/ryko.html

[4] http://www.clir.org/pubs/abstract/pub96abst.html

[5] see for example http://www.solarius-inc.com/index.html

[6] see for example http://www.veeco.com/technologies/stylus.asp

[7] See for example: http://www.elpj.com/

[8] See for example http://members01.chello.se/christer.hamp/phono/petrov.html and http://members01.chello.se/christer.hamp/phono/poliak.html

[9] Fruhwirth, R., Regler, M., Bock, R.K., Grote, H., and Notz, D., **Data Analysis Techniques for High Energy Physics, 2nd Edition,** Cambridge University Press, August 2000, ISBN-0521635489

[10] Alston, M., Franck, J.V., and Kerth, L.T., in **Bubble and Spark Chambers Vol.2**, p51-140, Shutt,R.P.,editor, Academic Press, 1967, LOC Cat number 6719642, and Hough, P.V.C., ibid, p141-194

[11] Aoki, S. et al, **Nucl.Instrum.Meth**.A473:192-196,2001

[12] EIA Specification RS-211-Rev D

[13] Read, Oliver, **The Recording and Reproduction of Sound**, H.Sams, 1953

[14] Langford-Smith, F., editor, **Radiotron Designer's Handbook**, Fourth Edition. The Wireless Press, 1953.

[15] Olson, H.F., **Acoustical Engineering**, van Nostrand, 1957, LOC Cat number 578143

[16] See for example: http://sound.westhost.com/project91.htm

[17] Rayner, F.J.W., Vaseghi, S.V., and Stickells, L., in **Archiving the Audio-Visual Heritage**, FIAF Symposium, May 20-22, 1987, p109

[18] See for example:
http://www.edmundoptics.com/TechSupport/electronic_imaging.cfm

[19] Optical Gauging Products, Inc. Web address http://www.ogpnet.com .

[20] MeasureMind Plus V10 is a product of Optical Gauging Products, ibid

[21] See for example:
http://ccrmawww.stanford.edu/CCRMA/Courses/422/projects/WaveFormat/ WAV for definition of the WAV file format. The *makewav* program by Don Cross converts PCM data as a list of ASCII integers into the WAV file format. It was available at http://www.intersrv.com/~dcross/wavio.html

[22] Decca Recording N 76422, "**Goodnight Irene**", by Gordon Jenkins and His Orchestra and The Weavers,

[23] Ramses II variable speed turntable from Esoteric Sound Inc., http://www.esotericsound.com/home.htm

[24] The cartridge used was part number M35x and the stylus was part number N78S both manufactured by Shure Inc., http://www.shure.com/catalog.html

[25] The preamplifier used was Model 40-630 Solid State Phono Pre-amplifier, manufactured by MCM. It was measured to have the standard RIAA frequency response.

[26] The sound card used was the Audiophile 2496 manufactured by M-Audio Inc., http://www.m-audio.net/products/m-audio/audiophile.php

[27] The audio editing software used was Sound Forge V6.0, a product of The Sonic Foundry Inc., http://www.sonicfoundry.com/

[28] The correct equalization curve to apply is probably the Decca78 version from reference [16]. The specific values for that were 3.4 KHz for the treble -3 dB point and 150 Hz for the bass +3 dB point. For the RIAA curve the values are 2.12 KHz and 500 Hz for treble and bass respectively. This is somewhat more restrictive at high frequency but it was judged most important to compare with the same curves. Work is in progress to obtain a pre-amplifier with the Decca78 characteristic.

[29] **The Weavers – The Best of the Decca Years**, MCA Records, 1996, MCAD-11465, track 3, “Goodnight Irene”.

[30] Marty Wekser, private communication

[31] The audio clips can be accessed at http://www-cdf.lbl.gov/~av/

[32] Hubbard, J.R.,**LHC Trigger Design,** 1997 CERN School of Computing (CSC 97), Pruhonice, Prague, Czech Republic, 17 Aug – 30 Aug 1997, Published in *Prague 1997, Computing* 117-141

[33] Clarke, P., **Nucl. Instrum. Meth.A368:175-178,1995**

[34] Cavaglieri, S., Johnsen, O., and Bapst, F., **Proc of AES 20th International Conference**, Budapest, Hungary 2001, Oct 5-7.

[35] http://www.cs.huji.ac.il/~springer/