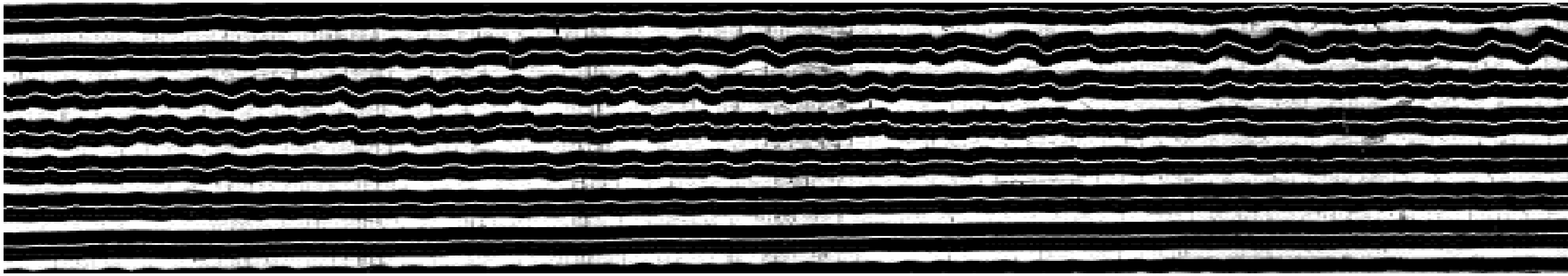
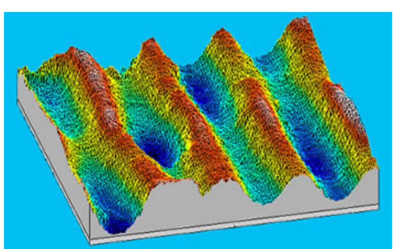


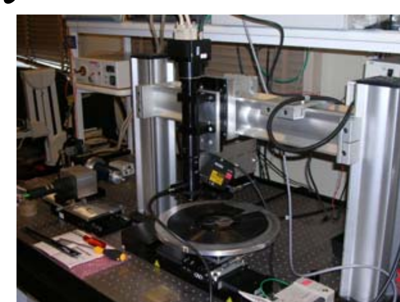
Using Images
to
Reconstruct Sound Recordings: Progress and Prospects

Update on Collaboration with the Library of Congress



Carl Haber

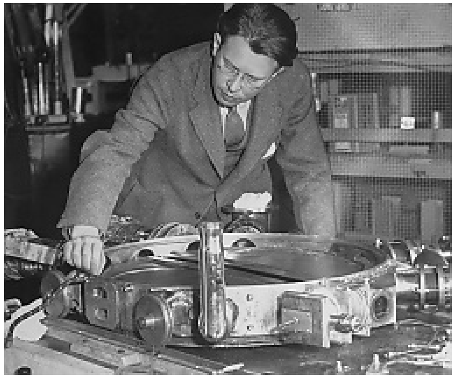
Lawrence Berkeley National Lab



Lawrence Berkeley National Lab

www.lbl.gov

- Founded in 1931 by E.O.Lawrence
- Oldest of US National Labs
- Operated by the University of California for the US DoE
- 4000 Staff, 800 Students, 2000 Guests
- 14 Research Divisions including
  - Physics, Nuclear Science
  - Materials, Chemical Science
  - Life Sciences, Physical Bioscience
  - Energy and Environment, Earth
  - Computing
- Major user facilities-
  - Advanced Light Source
  - Nat. Center for Electron Microscopy
  - Nat. Energy Research Super Computer Center



Outline

- Introduction
- Summary of method (mostly a repeat)*,**
- The I.R.E.N.E. Project – a fast disc scanner, now funded by NEH
- Research on 3D methods for cylinders & discs
- Public outreach and prospects
- Conclusions

* V.Fadeyev & C. Haber, **J. Audio Eng. Soc**., vol. 51, no.12, pp.1172-1185 (2003 Dec.).

** V. Fadeyev et al, **J. Audio Eng. Soc**., vol. 53, no.6, (2005 June). (to appear)

Collaboration and Support

Vitaliy Fadeyev, Carl Haber, Jian Jin, Stephen Wu
Lawrence Berkeley National Lab
Christian Maul, John W. McBride
Taicaan Technology, U.K., University of Southampton, U.K
Mitch Golden
Peter Alyea, Larry Applebaum, Elmer Eusman, Dianne van der Reyden
The Library of Congress
Mark Roosa
Pepperdine University
Sam Brylawski
University of California
Bill Klinger
ARSC
George Horn
Fantasy Records, Berkeley

Support from the Library of Congress, DoE, NEH, The Mellon Foundation

Corporate equipment donations, discounts, loan, and services from Aerotech, Archeophone, JH-Technology, Keyence, Microphotonics, National Instruments, Navitar, Newport, STIL, Uniforce, Veeco, Vidipax, Visicon

Introduction

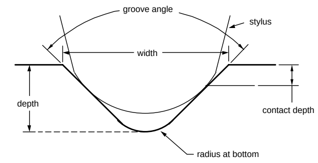
- We are developing methods of optically recovering mechanical sound recordings **without contact** to the medium
- Address concerns of the preservation, archival, and research communities:
  - Preservation: The reconstruction of delicate or damaged media
  - Access: Mass digitization of diverse media

- The approach evolved naturally out of methods of optical metrology, pattern recognition, and image processing.
- First shown at the LC in July 2003.

Traditional Contact Playback

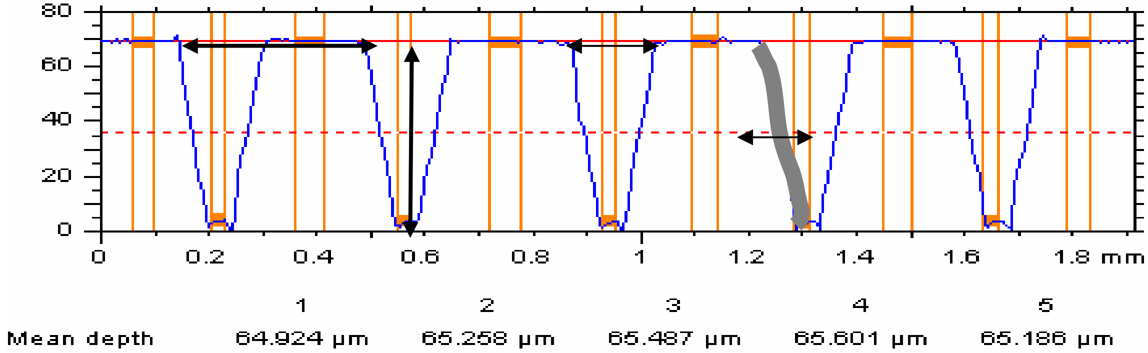
Bulky stylus riding in a narrow groove => Issues with
- tracking
- condition of the groove
    - debris and contamination
    - wear

Transcription process: requires trained manpower or supervision.

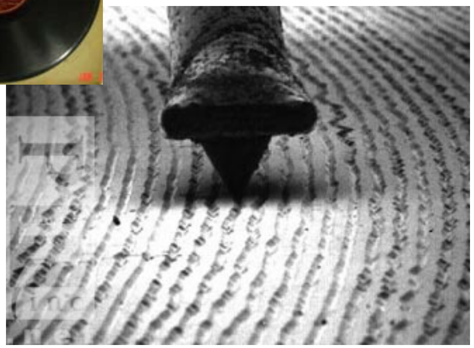


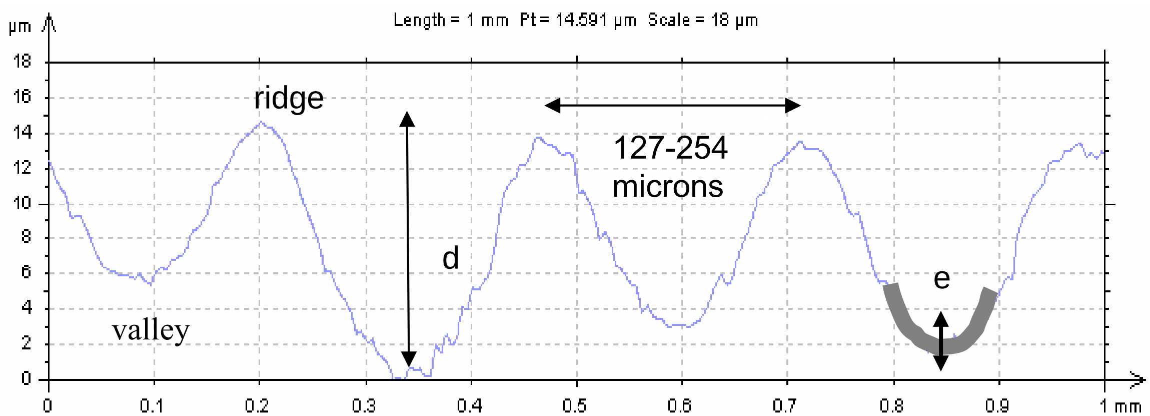
400 µm

150 µm

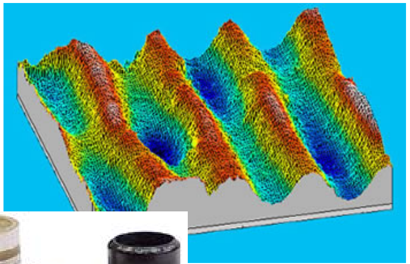


Disc surface





Cylinder surface



Measurement Challenge

Parameter	78 rpm, 10 inch	Cylinder
Cut	Lateral	Vertical
Area containing audio data	38600 mm^2	16200 mm^2
Total length of groove	152 meters	64-128 meters
Max groove amplitude (microns)	100 - 125	~10
Groove depth (microns)	80 fixed	+/- 10 varies
Groove displacement @noise level	1.6 - 0.16 microns	< 1 microns

Need to measure sub-micron features over entire surface of record

Non-Contact Digital Imaging

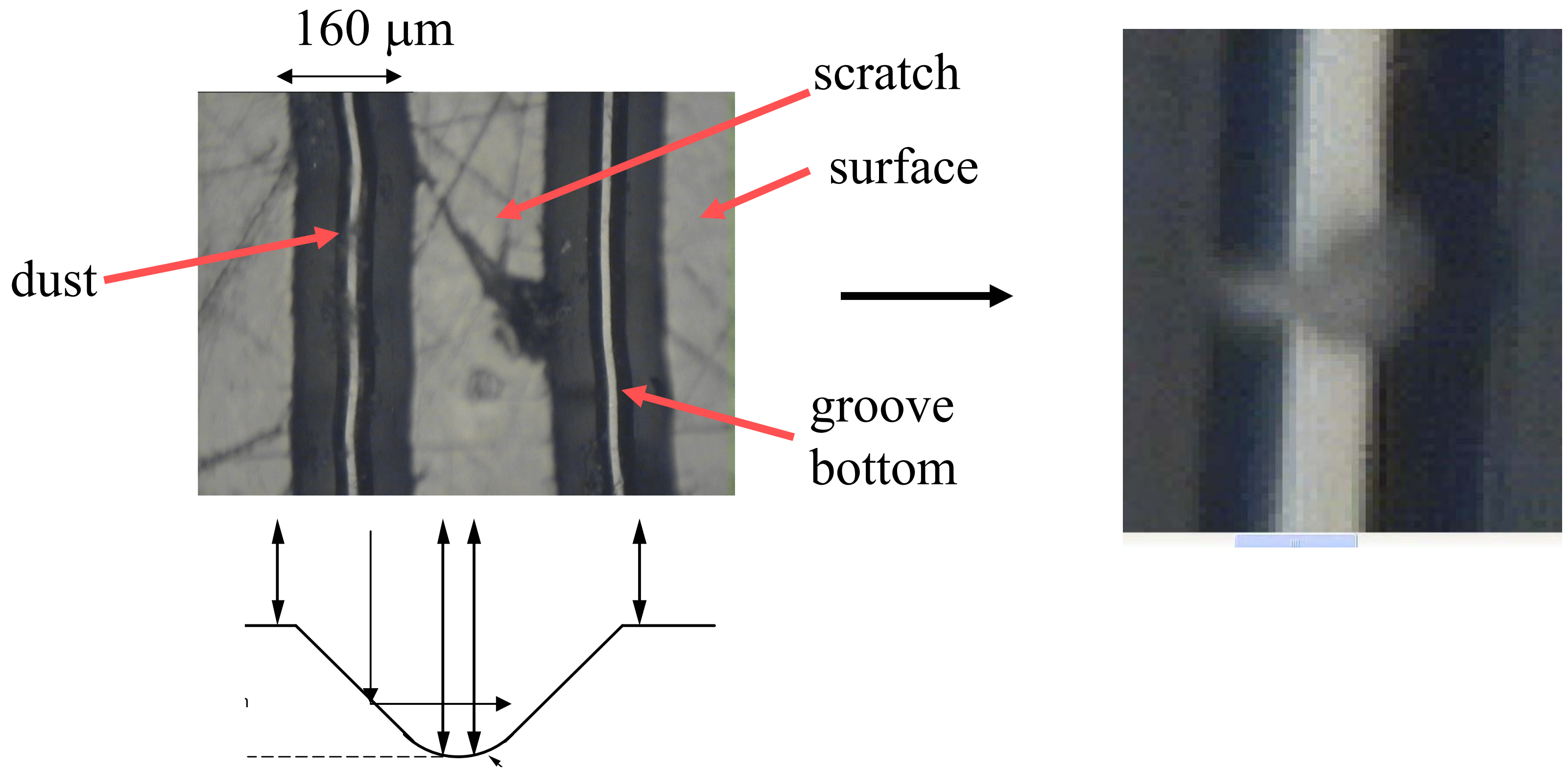
- Protects samples from further damage
- Repair existing damage through "touch-up"
- Offload aspects of restoration to automated software

A "smart" copying machine for records

The Method

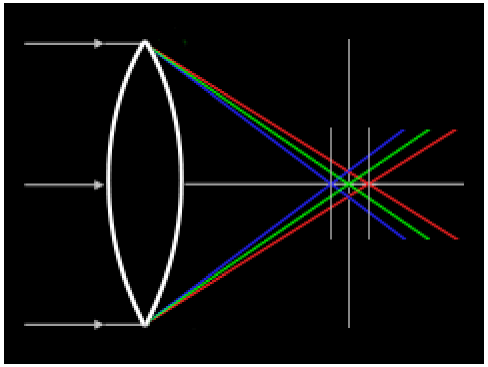
- Digitally image the surface
- Cover with sequential views or grid.
- Stitched together: surface map
- Process image to remove defects
- Analyze shape to model stylus motion.
- Sample at standard frequency
- Convert to digital sound format.
- Real time playback is not required

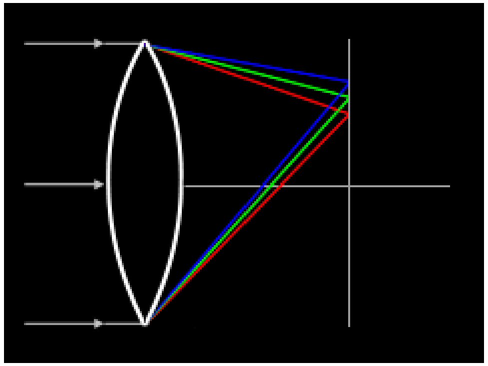
2D Imaging: Electronic Camera



- Suitable for disc with lateral groove
- Require 1 pixel = ~ 1 micron on the disc surface

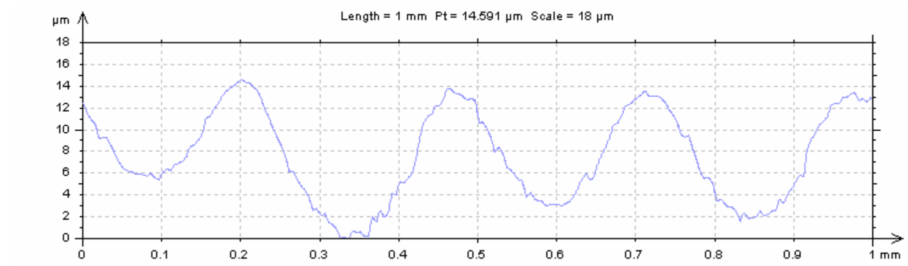
Chromatic Aberration

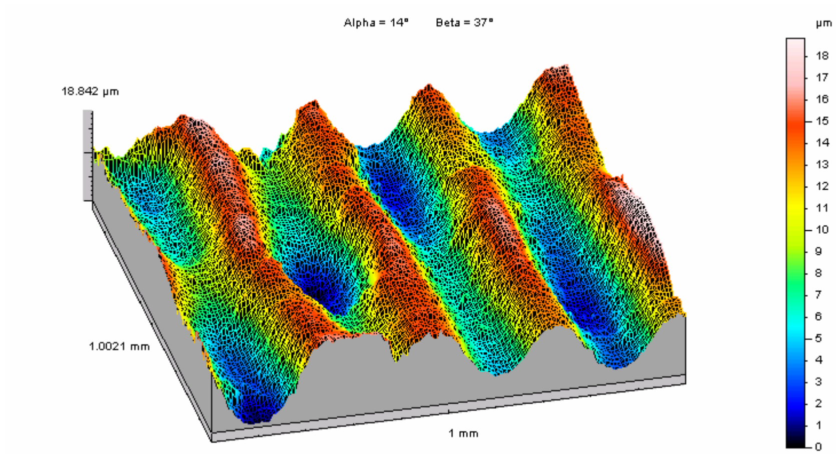


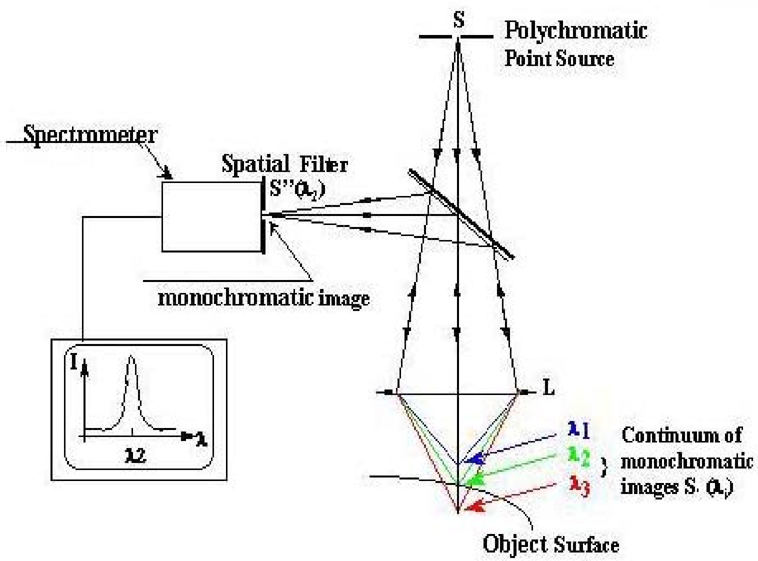


3D Imaging: Confocal Scanning Probe

Required for cylinder with vertical groove modulation.







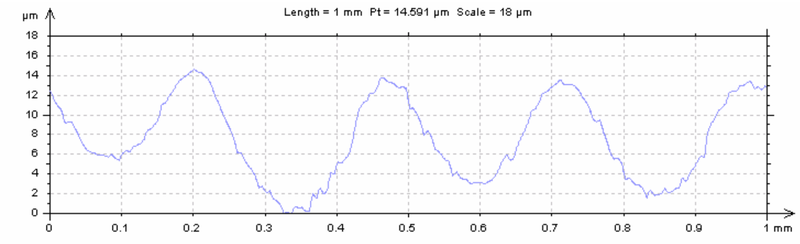
Surface of an Edison cylinder

Reconstruct the surface from a large set of points

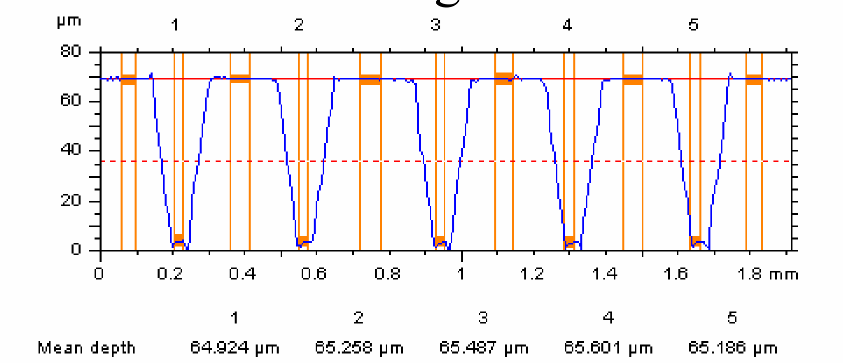
Speed and Data

- 2D scans for lateral discs (IRENE)
  - Line scan camera: ~5-15 min/scan for 10" 78 rpm disc
  - 50 Mb / 1 s of raw images
  - 1.5 Mb / 1s of processed images
  - 88 Kb / 1s audio (4416)
- 3D scans for vertical media
  - Depends upon grid, probe rate
  - 12 KHz time sampling: 3-10 hours
  - 96 KHz time sampling: 24-80 hours
  - Additional factors of 2-4 may be available soon
- 3D for deep groove lateral discs
  - Much slower probe rates are probably required

Vertical groove



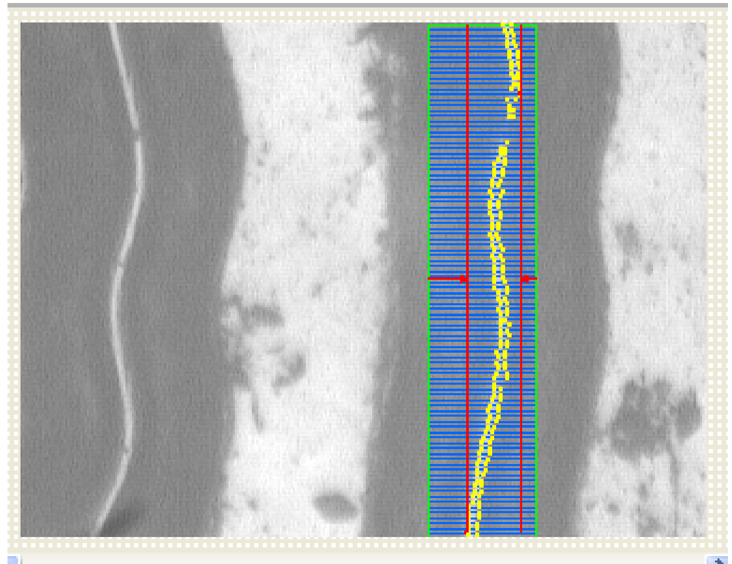
Lateral groove



Key 3D issues are slope and depth

Image Processing

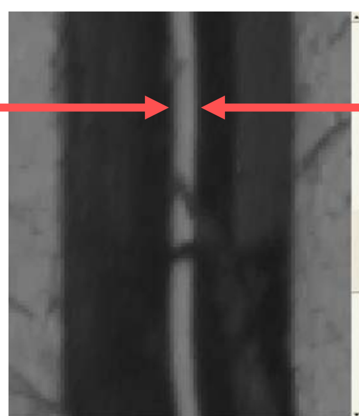
Intensity

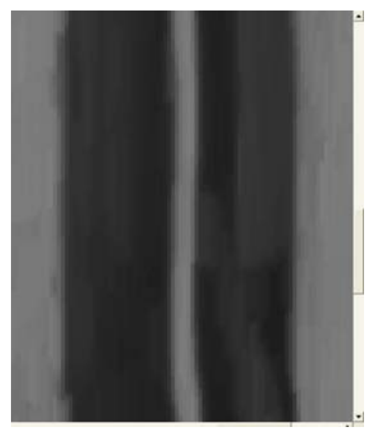


Edge finding

Knowledge of groove geometry provides a powerful constraint for rejecting debris and damage

Groove Geometry constraint





dilation

Signal Analysis

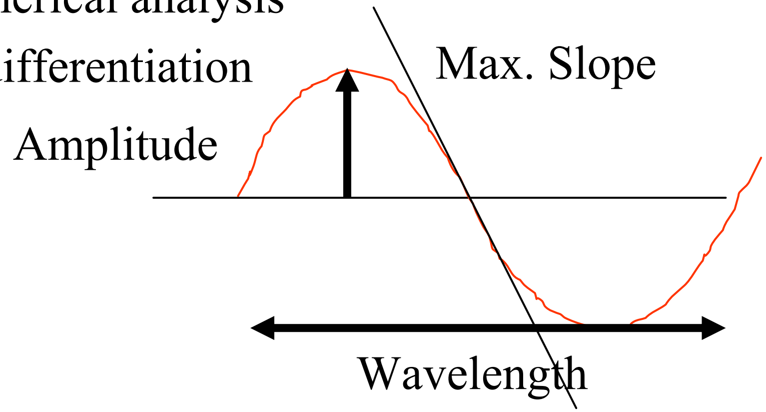
- For recording and playback, signal is proportional to stylus velocity

  "electrical": magnetic induction

  "acoustic": plane wave approximation, air pressure and velocity are proportional and in-phase

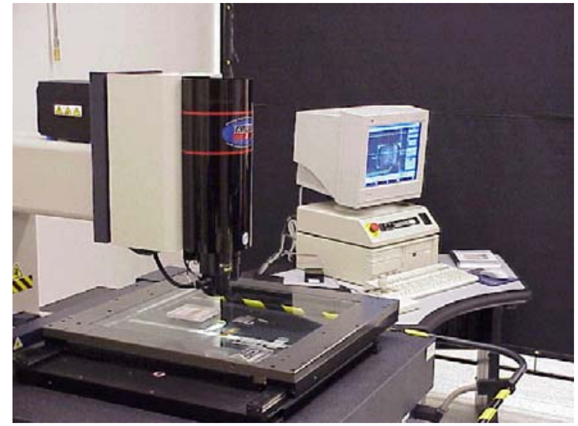
- Electrical recordings are (deliberately) mediated by equalization scheme to attenuate low frequencies and boost high frequencies
- Acoustic recordings are (naturally) mediated by the frequency response of horns and diaphragms.
- Potential to improve fidelity with modeling of acoustic component response
- Groove data is in digital form, numerical analysis
- Determine velocity by numerical differentiation

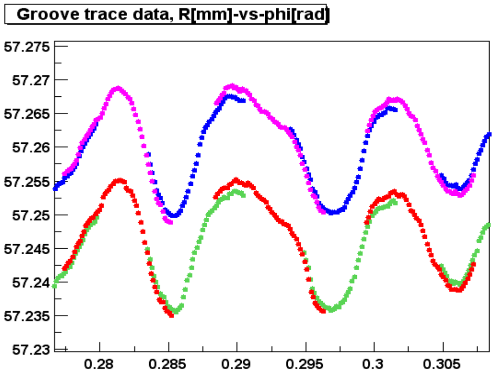
$$A_p = \frac{v_p}{2\pi f}$$



Original Test of Concept (2002)

- Use available tools
- Optical metrology system "SmartScope" manufactured by Optical Gauging Products.
- Image acquisition with pattern recognition and analysis.
- Programmed to scan groove, report, and process data (offline).
- Study of 78 rpm shellac disc ~1950
- Very slow: 40 minutes / 1s of audio

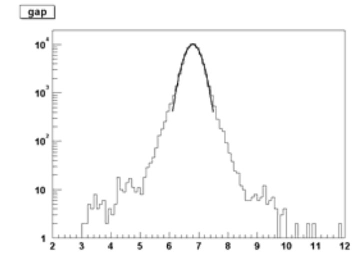




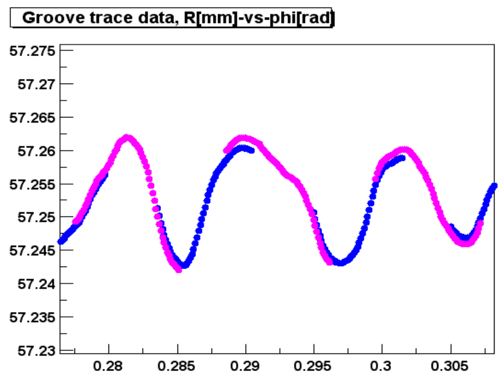
Raw measurement
of groove bottom edges

ΔR

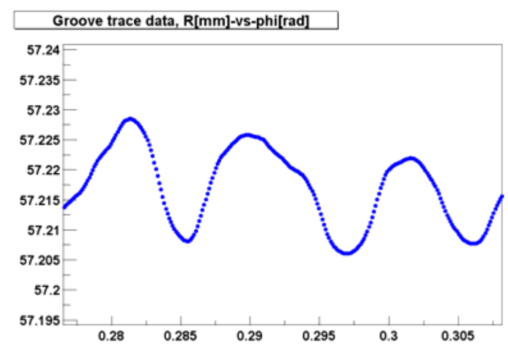
ΔR distribution



Width across
groove bottom



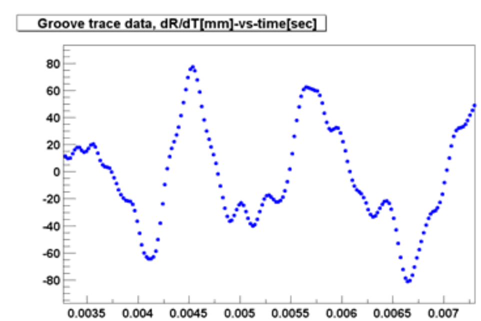
Averaged and filtered
for known width
ΔR<cut



Frames aligned

Measurement spacing along
time axis ~ 66 KHz

Stylus velocity by
numerical differentiation
4th order polynomial fit of
15 points about each sample



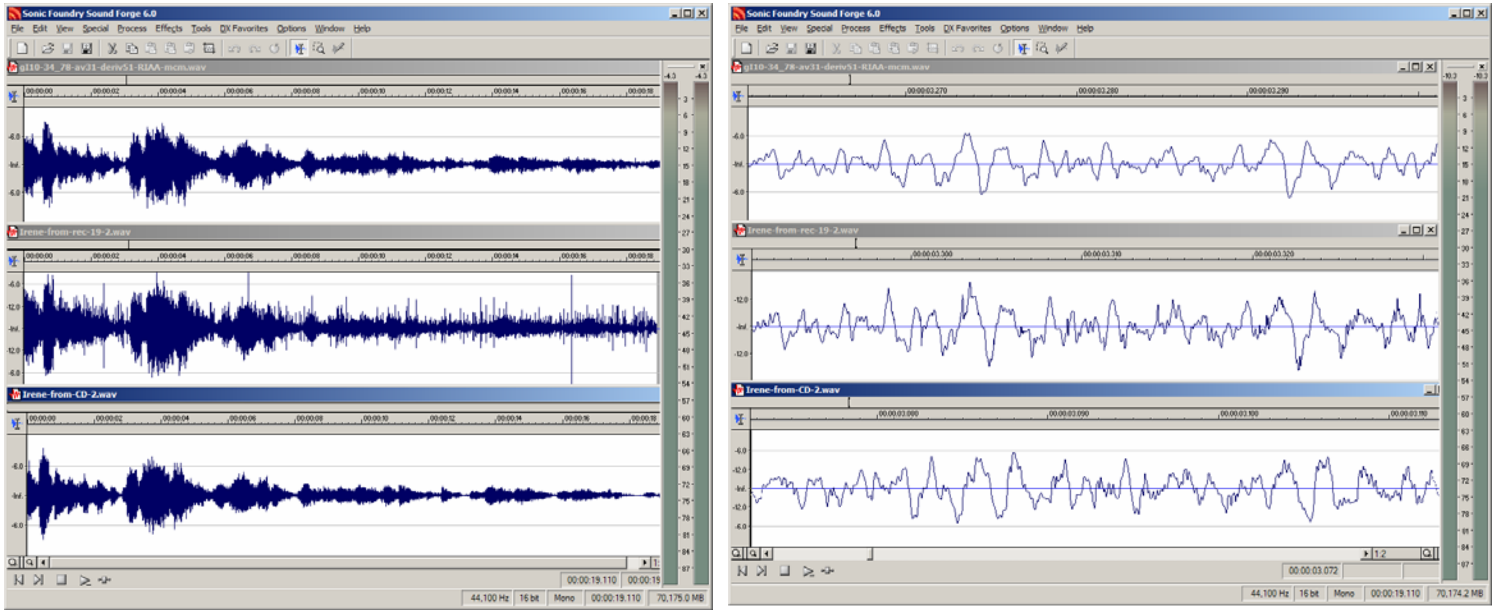
Waveform comparison

19.1 seconds   40 ms

optical

stylus

CD

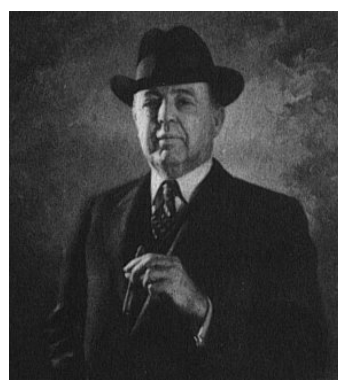


- Clear reduction in “clicks and pops”
- Similarity of fine waveform structure

Sound Comparison

"Goodnight Irene" by H. Ledbetter (Leadbelly) and J.Lomax, performed by The Weavers with Gordon Jenkins and His Orchestra ~1950

- Sound from the CD of ***re-mastered tape***.
- Sound from the ***mechanical (stylus)*** readout.
- Sound from the ***optical*** readout.
- optical + commercial noise reduction



LC Directions

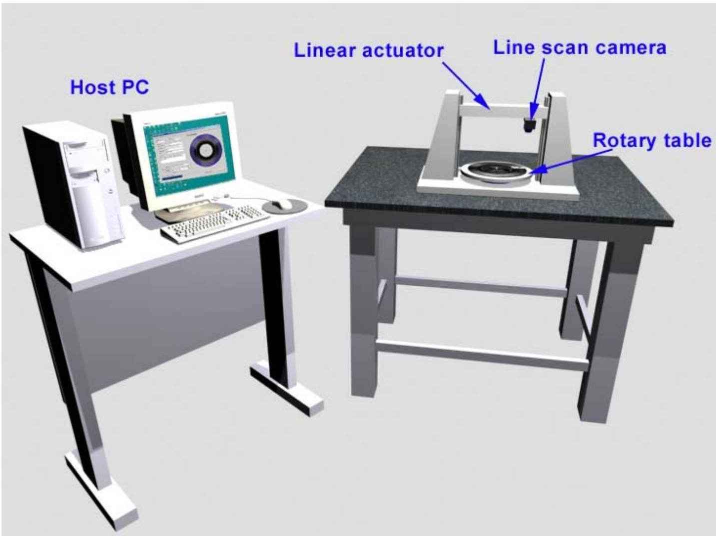
1. The 2D test was promising, can you make a machine to run near real-time on discs? Could it address mass digitization needs? What about sound quality?
   - IRENE proposal
2. A research program to further the 3D technology.
   - Underway with support from LC, Mellon

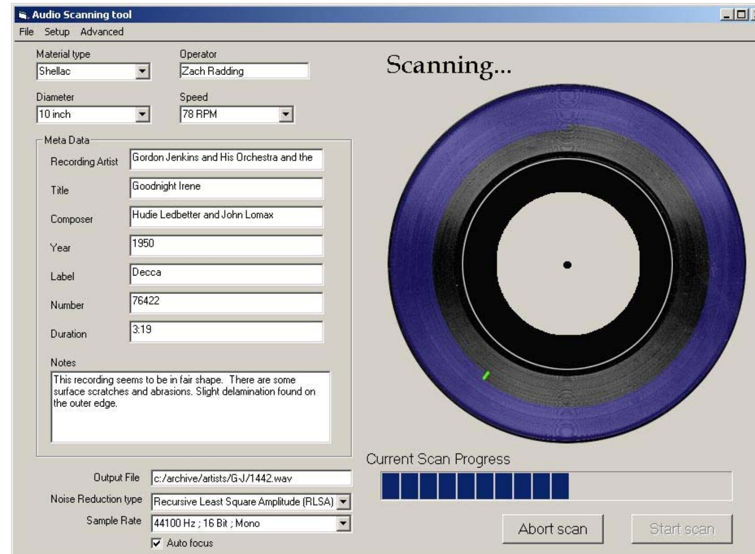
Design of a 2D Machine

- 2003: LBNL supports design study for 2D machine
- LC: specs and media samples.
- Design, cost, schedule pass internal reviews at LBNL.
- Proposed to NEH 7/04, approved 1/05, de-scoped
- 2005: testing of real components in preparation of full engineering and coding process – this presentation
- Construction now – 3/06

I.R.E.N.E.

Image, **R**econstruct, **E**rase **N**oise, **E**tc





- ~1 year development and construction
- Projected scan time 5-15 minutes
- Provide statistical measures of media condition
- Production-like machine and test-bed for future development

Basic Features and Goals

- 2D approach: image groove bottom and/or top.
- Emphasize throughput.
- Encompass as much variation in media as possible.
- Handle broken discs.
- Facility to (temporarily) flatten flexible media (Memovox)
- User friendly interface.
- Commercial off-the-shelf components.
- Provide a test bed for the mass digitization application.
- Project to be managed through LBNL Engineering Division, Dr. Jian Jin is PI

Testing and Validation Phase

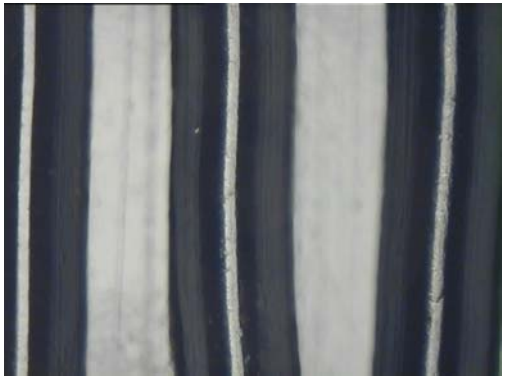
- LC is to receive machine in early 2006
- Machine to be tested on a significant sample of media by LC staff
- Results to be documented and disseminated
- Possibility of a follow-on stage
  – Upgrades to the machine in software (easy) or hardware
  – Expanded media study

Advisory Panel

- Representation from technical or audio experts
  - Adrian Cosentini (Vidipax)
  - Vitaliy Fadeyev (LBNL)
  - George Horn (Fantasy Studios)
  - Charles Mayn (NARA)
- Representation from “user community”
  - Sam Brylawski (UC)
  - Jake Horniak (Smithsonian)
  - Mark Roosa (Pepperdine)
  - Anthony Seeger (UCLA)
  - Abby Smith (CLIR)
  - Sarah Stauderman (Smithsonian)

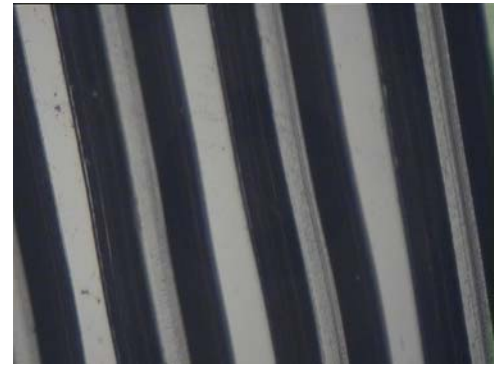
Media Condition Survey on LC Samples

good 65%



poor 10%

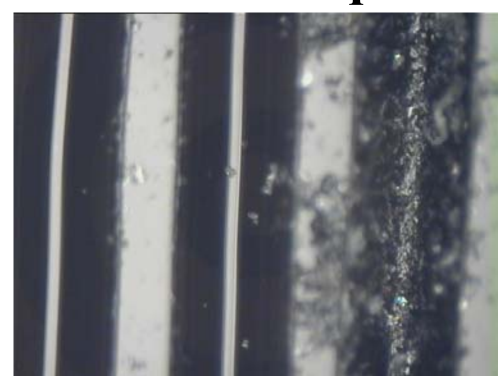
good lacquer



fair 25%

fair

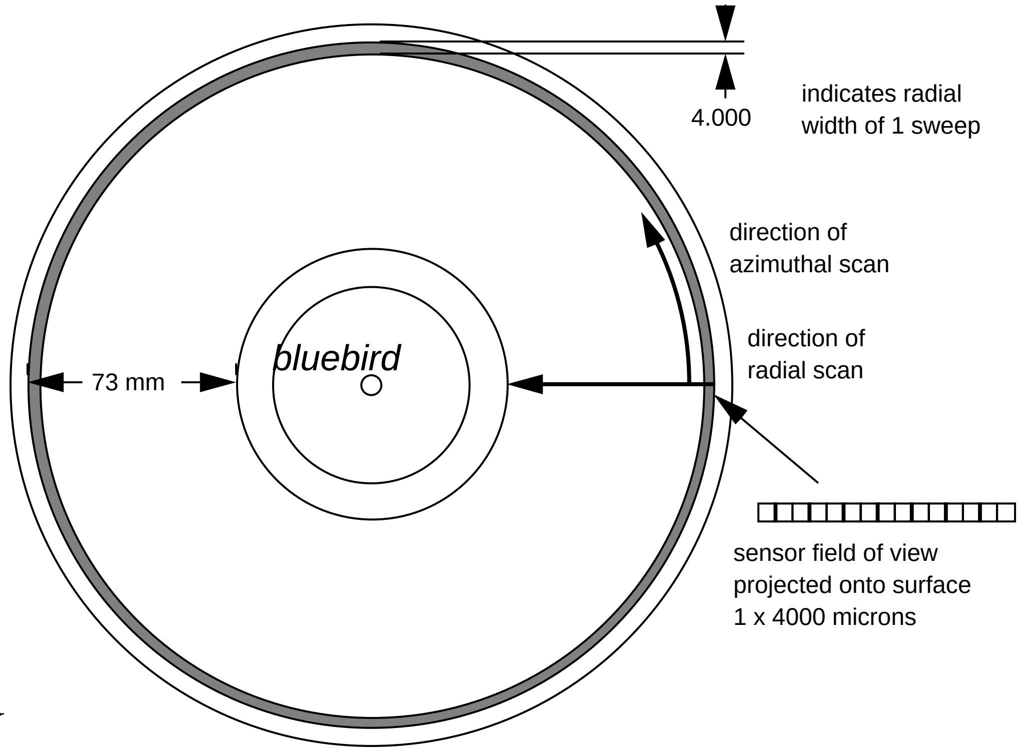
exudated lacquer



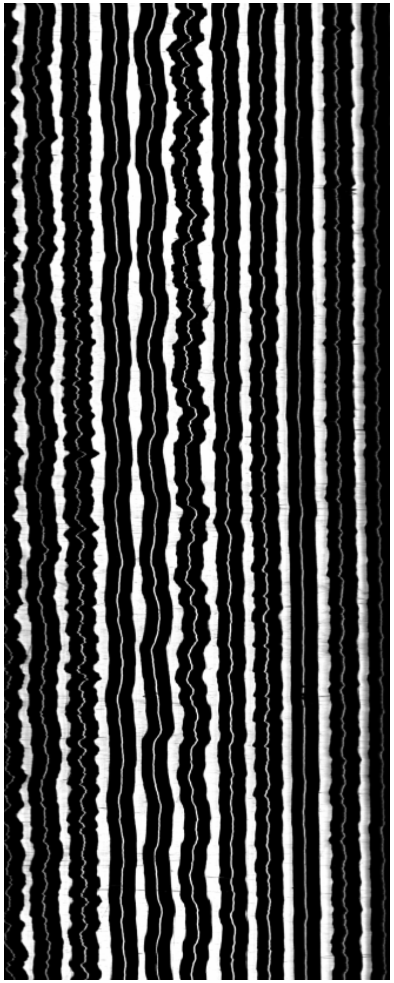
Multiple edges

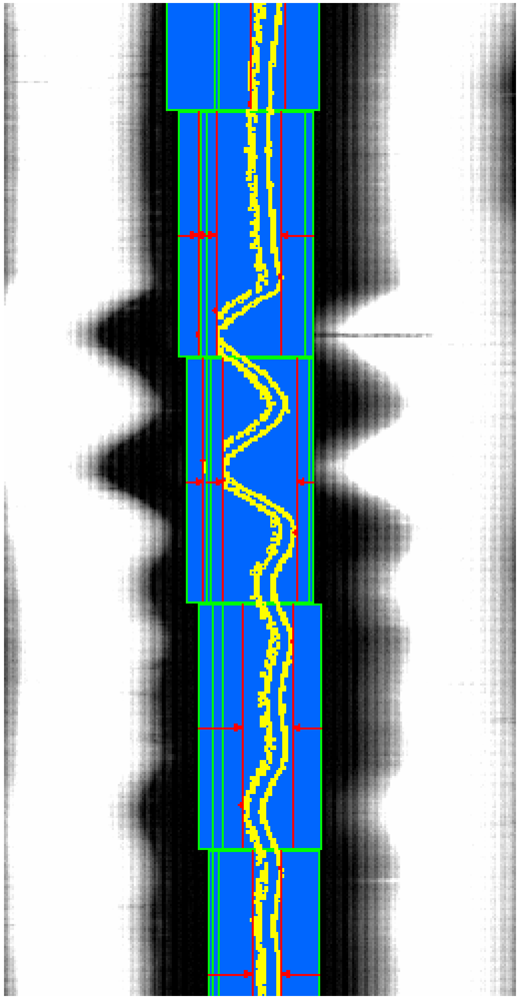
Rough groove bottom

- 4000 pixel, 18 KHz line scan sensor
- Magnify to 1 pixel = 1 μm
- 7.6 x 10^5 lines/outer ring
  - 390 KHz sampling
- Time/ring = 40 seconds
- 73 mm / 4 mm = 19 rings
- 19 x 40 sec = 13 minutes
- Reduce with variable speed on inner rings: 9 minutes
- Scan time decreases linearly with sampling!!!.



Based upon 10 inch, 78 rpm geometry

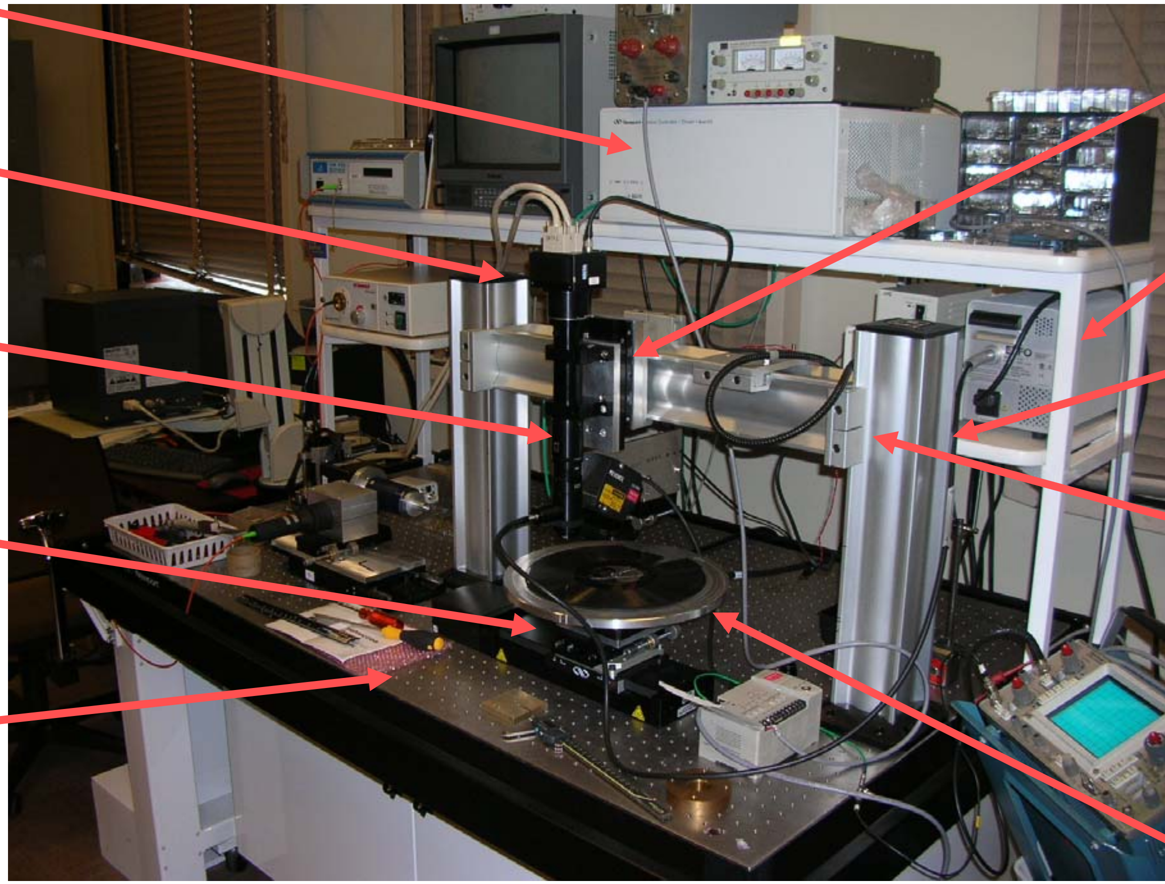




Performance

Feature	Targets	Best Effort	Upgrade
Coarse grooves, track and measure groove bottom or top	X		
Measure additional features on groove wall		X	X
Use of coaxial lighting	X		
Use of alternative lighting approach		X	X
Reconstruct GOOD based on LOC sample set (Figure 7)	X		
Reconstruct FAIR based on LOC sample set (Figure 8)		X	X
Reconstruct POOR based on LOC sample set (Figure 9)			?
Scan time of 6-15 minutes on flat samples	X	X	
Processing time of 3-5 minutes	X	X	
Scan time on warped samples – as fast as practical		X	
Reconstruct Memovox or other highly warped items		X	
Reconstruct broken records, good quality debris	X		
Reconstruct broken records, lower quality debris		X	X
Exudated lacquers – well cleaned (Figure 11a left)	X		
Exudated lacquers–moderate cleaning (Figure 11a right,b)		X	
Available sampling rates < 200 KHz	X		
Available sampling rates > 200 KHz		X	
Reconstruct stampers	X		
Data and statistical results on quality of scan and sample	X		

IRENE Test Platform



Stage motion controller

Line scan camera

Main lens

Motion stages

Vibration isolation table

Vertical stage for focus

Light sources

Fiber bundle

Support arch

Focus height sensor

Turntable and disc

Hardware Issues

- Linescan camera
  - Dalsa P2 4000 pixel, 18KHz, NI CameraLink interface
- Optics
  - Navitar 12x zoom with internal coaxial port – FOV issue
- Motion and interface
  - Newport and Aerotech products evaluated
- Focus
  - Keyence LKG series laser triangulation sensors
- Illumination
  - EXFO Omnicure, liquid light guide – exposure time issue
- Processing
  - Matrox RT Image processor to be evaluated – data flow and storage

Software Issues

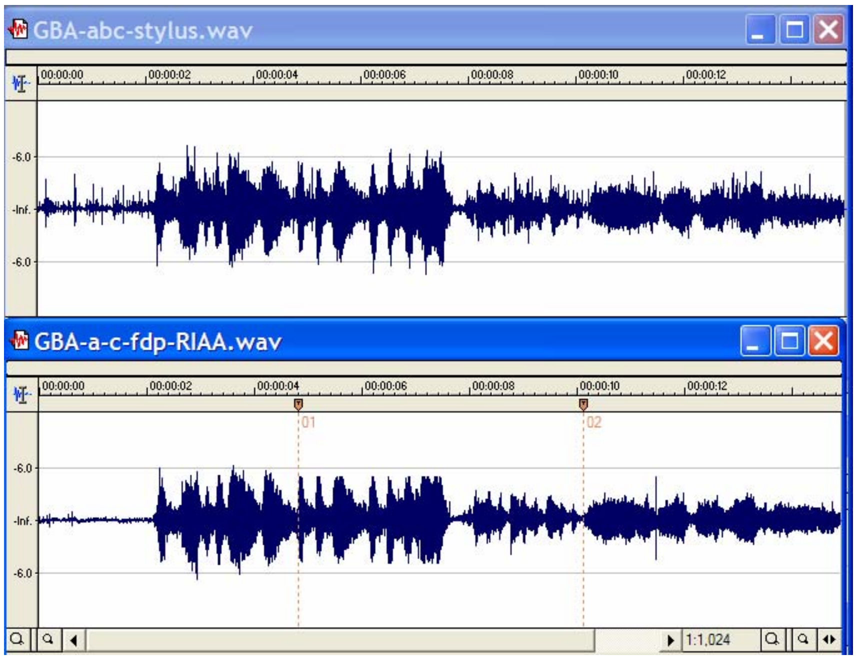
- Considerable software for control, data taking, and analysis developed for tests
  - Utilize National Instruments LabView system and their IMAQ/Vision software
  - Algorithms and requirements derived from this
- Need to understand appropriate software environment for actual machine
  - Bring programmers/engineers into the effort at this point

Tests of IRENE Design

- Preliminary results from prototype configuration
  - Used low intensity illumination
  - Camera 1K lines/s (use 18 K in real system) - 13 s scan/ 1 s audio
  - Raw time sampling was 104 KHz
  - RIAA curve applied for purpose of comparison with stylus
- Shellac 78 rpm discs
  1. Good condition
  2. "Good Night Irene" from original test (comparison)
  3. Very worn disc
  4. Distorted audio
- Acetates
  - 1950's studio music take
  - 1938 spoken word
- Photos of media array

Good Shellac: Waveform Comparison

Stylus



IRENE test platform

Sound Comparison

God Bless America

Composed by Irving Berlin, performed by Kate Smith
Victor release

Stylus version

Optical version using IRENE test platform

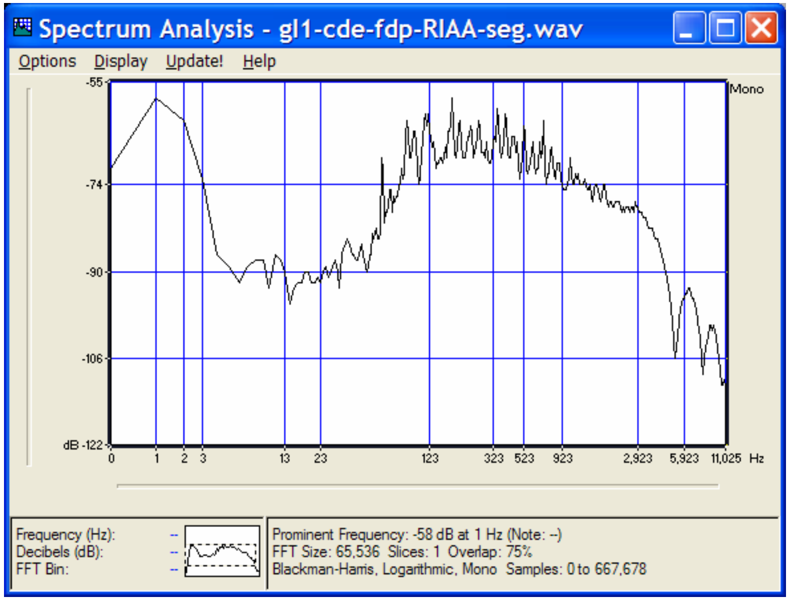
Test: Modulated Noise

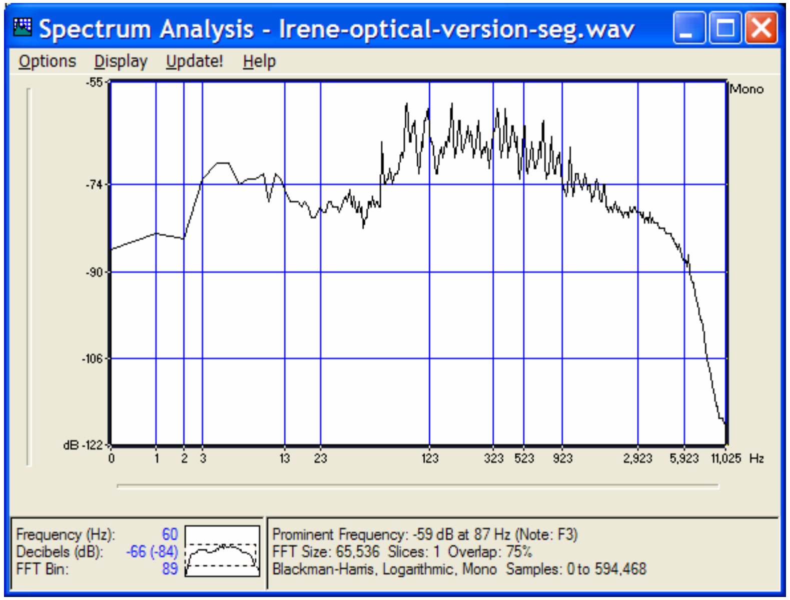
"Good Night Irene" disc used in original study

From original 2D optical study (40 min / 1 s of audio)

From IRENE test (13 s / 1 s of audio)

Spectra


Spectrum Analysis - gl1-cde-fdp-RIAA-seg.wav
Options Display Update! Help
Mono
-55
-74
-90
-106
dB -122
0 1 2 3 13 23 123 323 523 923 2,923 5,923 11,025 Hz
Frequency (Hz): --
Decibels (dB): --
FFT Bin: --
Prominent Frequency: -58 dB at 1 Hz (Note: --)
FFT Size: 65,536 Slices: 1 Overlap: 75%
Blackman-Harris, Logarithmic, Mono Samples: 0 to 667,678


Spectrum Analysis - Irene-optical-version-seg.wav
Options Display Update! Help
Mono
-55
-74
-90
-106
dB -122
0 1 2 3 13 23 123 323 523 923 2,923 5,923 11,025 Hz
Frequency (Hz): 60
Decibels (dB): -66 (-84)
FFT Bin: 89
Prominent Frequency: -59 dB at 87 Hz (Note: F3)
FFT Size: 65,536 Slices: 1 Overlap: 75%
Blackman-Harris, Logarithmic, Mono Samples: 0 to 594,468

Test: Defects etc.

<u>Dirty and worn</u>

When You and I Were Young, Maggie

Composer: Johnson and Butterfield

Performed by Charles Harrison

Victor 17474-B

Stylus version

IRENE test scan

<u>Some audio distortion</u>

Uchar Kupietz (folksong)

Performed by Vera Smirnova

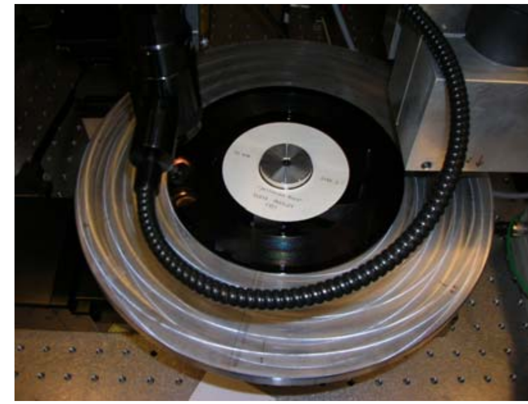
Columbia 20115-F

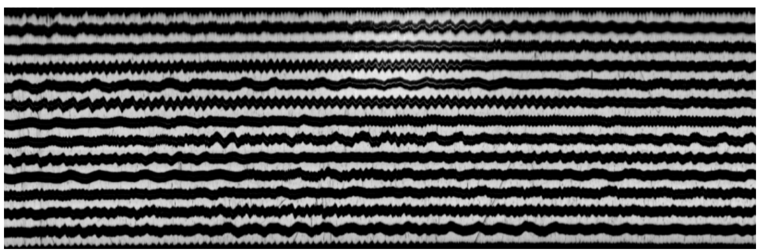
Stylus version

IRENE test scan

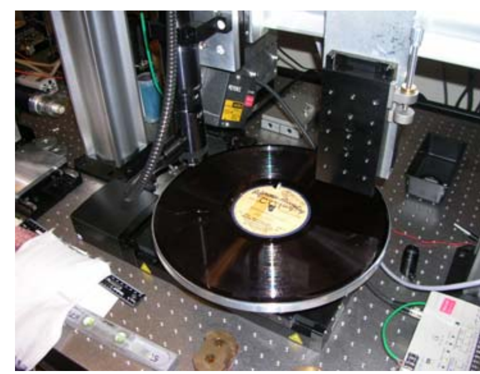
Acetates

45 rpm fine groove with
No groove bottom image
Labeled: “Jailhouse Rock,
RCA property”

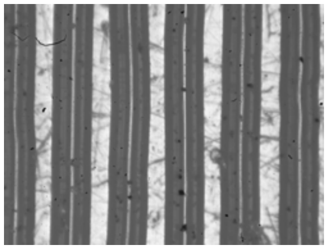


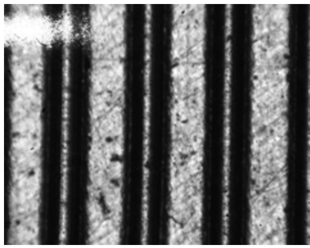


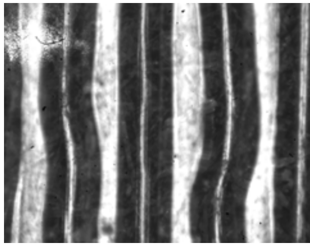
78 rpm lacquer on glass
Label: Howard Hughes,
Collier Award 1939

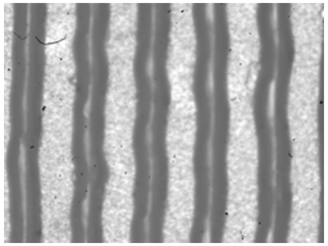


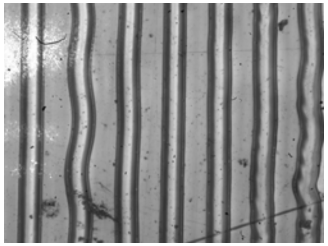
Unusual Media Survey: Images Reasonable











Thanks to
T. Aukofer

IRENE Summary

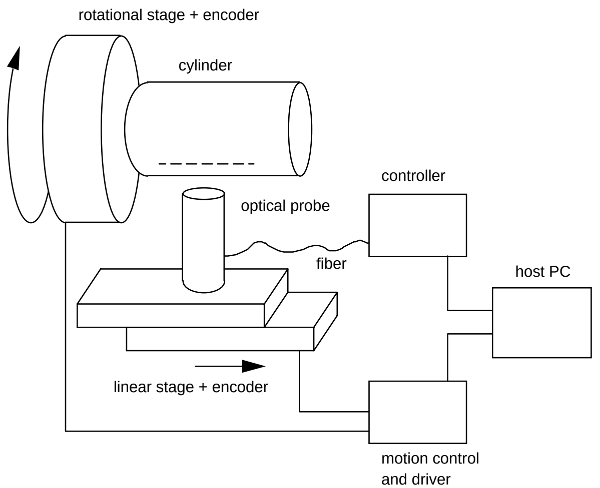
- Baseline hardware identified and tested.
- Projected scan time reasonable for a production-like machine.
- Software development will be the major task
- Would also provide a powerful test-bed for further development.
- Provides a new statistical view of disc media

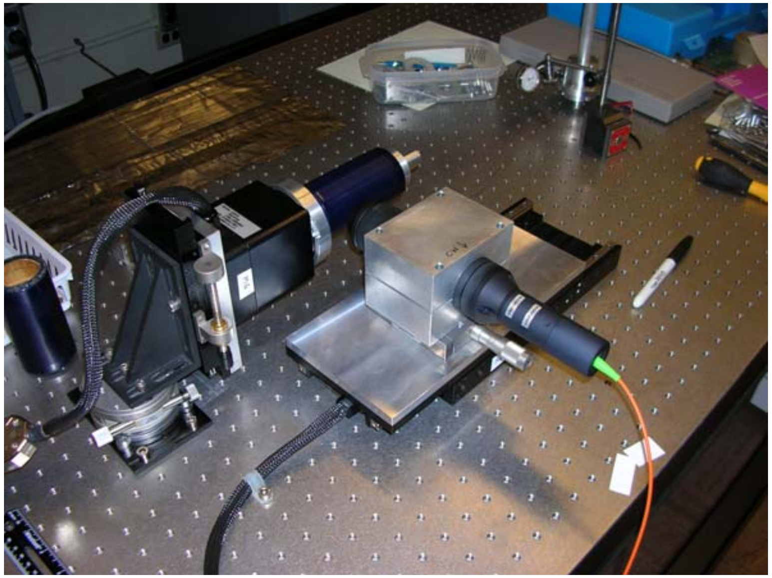
Test of 3D Scanning

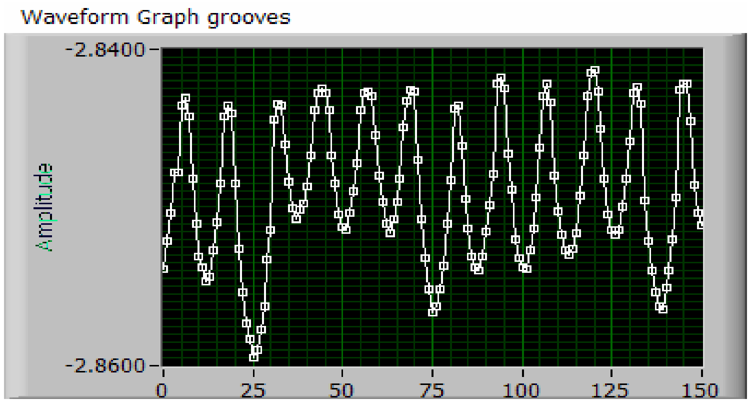
- Initial test in collaboration with Taicaan Technology and U of Southampton, UK
  - Presented at LC May 2004
  - To appear in JAES June 2005
- Scan speed was not emphasized, wanted to perform a proof-of-concept (~2 hr / 1 s audio)
- Faster 3D scanner now in operation at LBNL

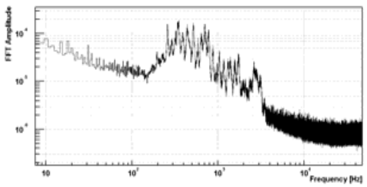
3D Study of an Edison Cylinder

Utilize confocal scanning probe at 300, 1000, 4000 Hz,
7.5 μm spot, 10 μm points
Angular increment = 0.08 - 0.01° = 12 - 96 KHz time sampling

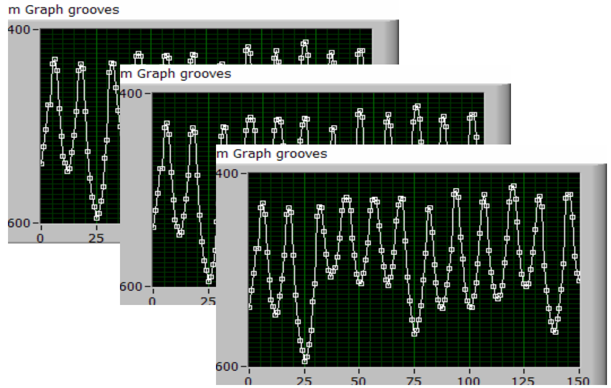




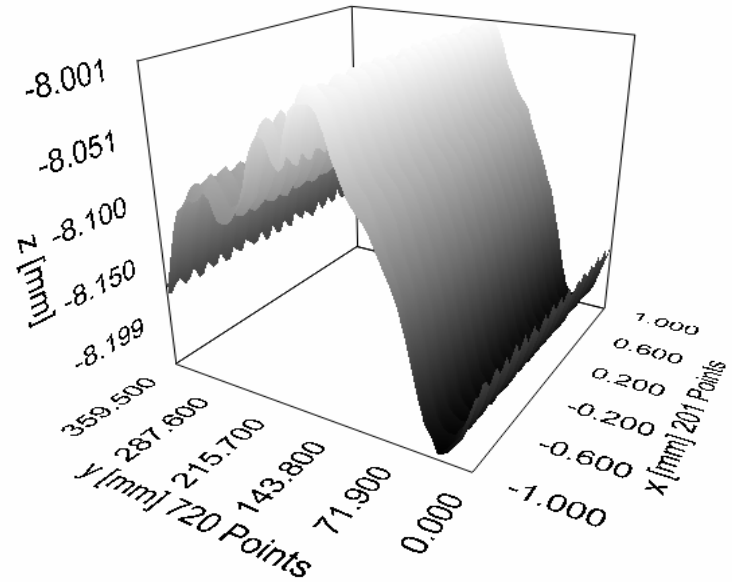


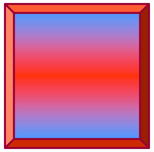


Sample at 96KHz to minimize effect of aliasing



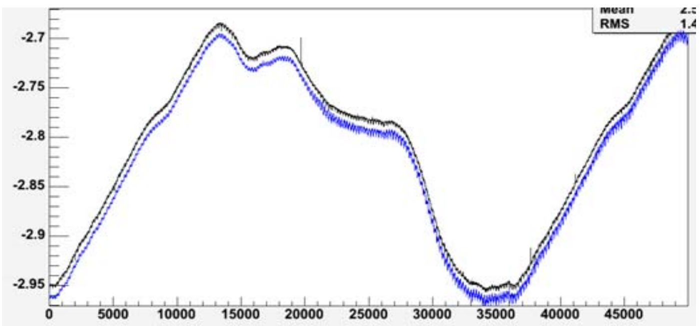
Sequential axial scans





Overall cylinder shape due to off-center, deformation, heard as low freq rumble

Subtract valleys from ridges to correct for overall shape



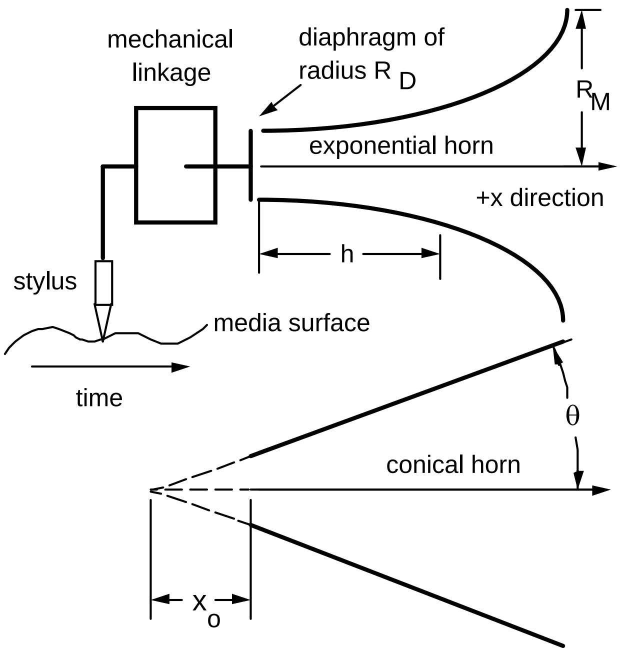
Ridge tops provide a useful, though imperfect, geometrical reference

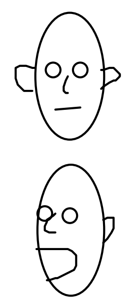
Sound Comparison

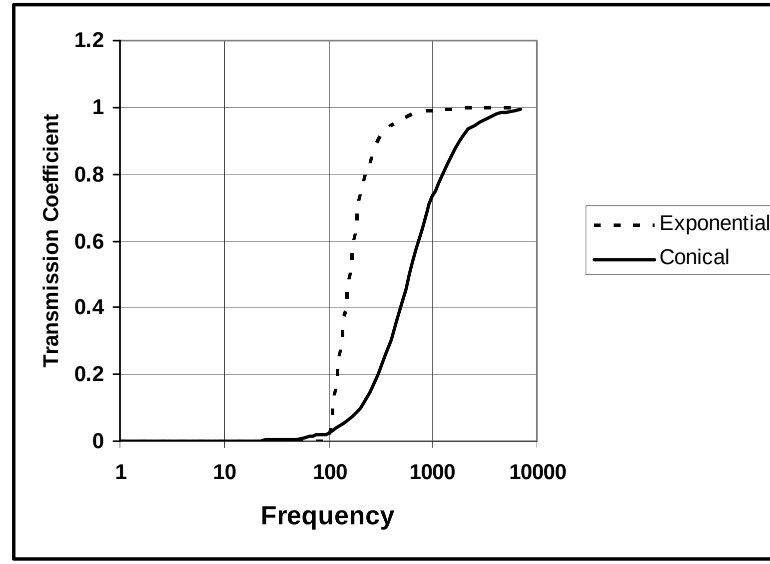
- The Holy City, composed by Stephen Adams,

The Edison and Skedden Mixed Quartet, Amberol 1601

- Stylus version - flat
- Optical version (1 KHz probe rate) - flat
- Optical version + commercial filter + EQ
- Optical version (4 KHz probe rate)

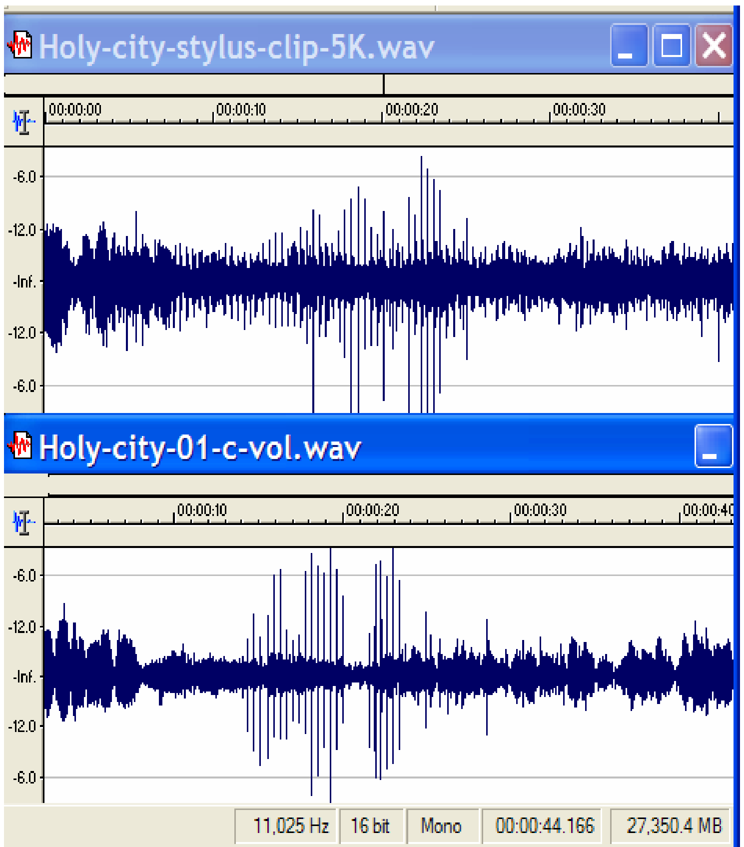


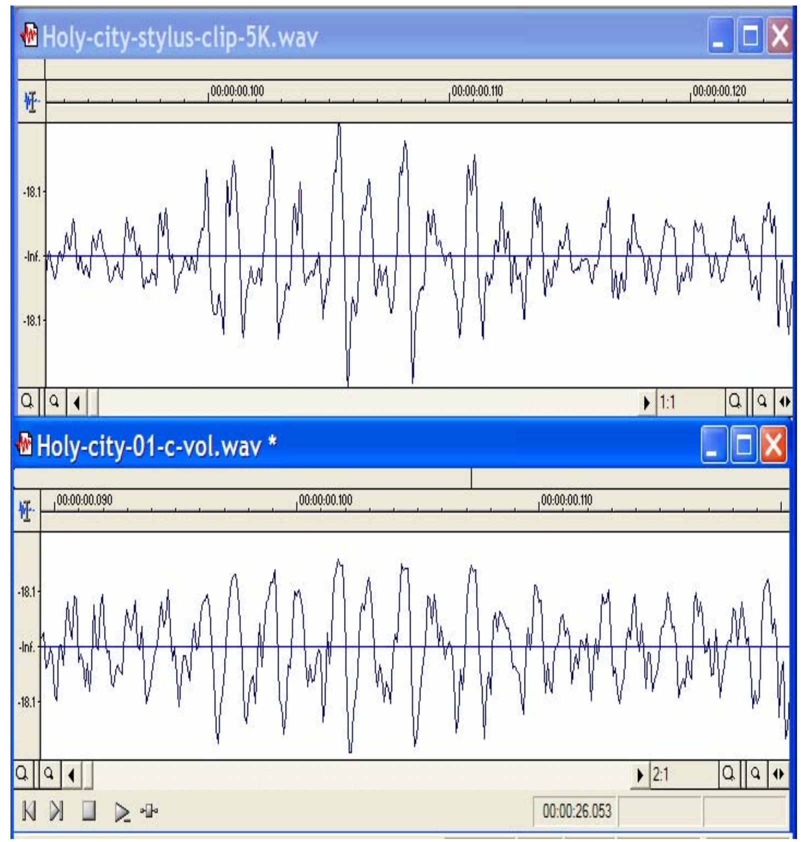


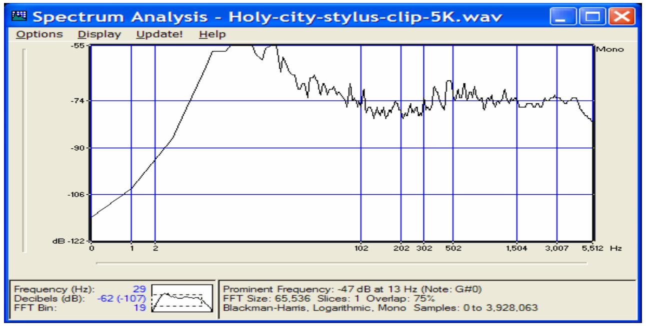


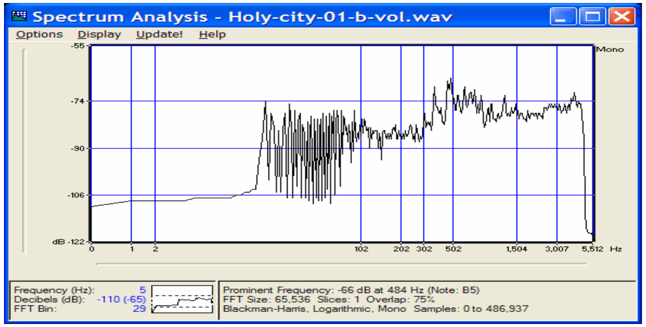
Response of horn and diaphragm at low frequency can modify response and deviations from "constant velocity" characteristic.

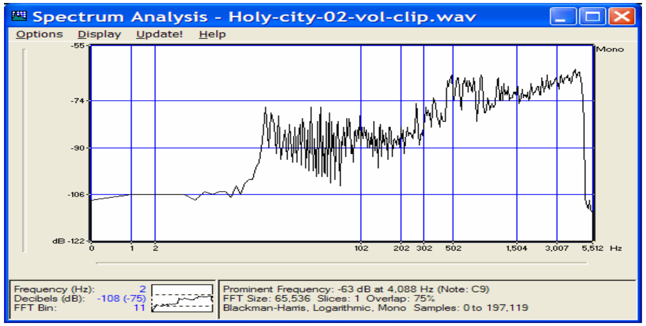
The Holy City: Waveforms






Spectrum Analysis - Holy-city-stylus-clip-5K.wav
Options Display Update! Help
Mono
-55
-74
-90
-106
dB -122
0 1 2 102 202 302 502 1,504 3,007 5,512 Hz
Frequency (Hz): 29
Decibels (dB): -62 (-107)
FFT Bin: 19
Prominent Frequency: -47 dB at 13 Hz (Note: G#0)
FFT Size: 65,536 Slices: 1 Overlap: 75%
Blackman-Harris, Logarithmic, Mono Samples: 0 to 3,928,063


Spectrum Analysis - Holy-city-01-b-vol.wav
Options Display Update! Help
Mono
-55
-74
-90
-106
dB -122
0 1 2 102 202 302 502 1,504 3,007 5,512 Hz
Frequency (Hz): 5
Decibels (dB): -110 (-65)
FFT Bin: 29
Prominent Frequency: -66 dB at 484 Hz (Note: B5)
FFT Size: 65,536 Slices: 1 Overlap: 75%
Blackman-Harris, Logarithmic, Mono Samples: 0 to 486,937

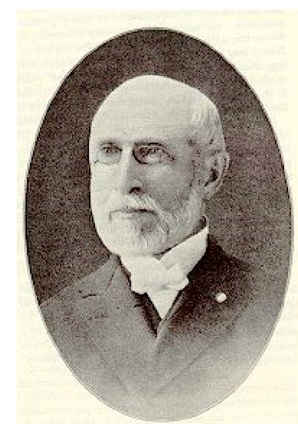

Spectrum Analysis - Holy-city-02-vol-clip.wav
Options Display Update! Help
Mono
-55
-74
-90
-106
dB -122
0 1 2 102 202 302 502 1,504 3,007 5,512 Hz
Frequency (Hz): 2
Decibels (dB): -108 (-75)
FFT Bin: 11
Prominent Frequency: -63 dB at 4,088 Hz (Note: C9)
FFT Size: 65,536 Slices: 1 Overlap: 75%
Blackman-Harris, Logarithmic, Mono Samples: 0 to 197,119

Sound Comparison

- “Just Before the Battle, Mother”, composed by George F. Root, performed by Will Oakland and Chorus 1909,

  1516 (..76; 4M-297-2) originally as Amberol #297 1909

- with stylus, flat equalization
- Optical version, flat equalization
- + commercial noise reduction + low frequency boost *



*thanks to George Horn, Fantasy Records, Berkeley

2nd Sound Comparison

- Hawaiian Guitar Duet

  Louise and Ferrara, Edison Blue Amberol 3065

- Stylus version
- Optical version (1 KHz probe rate)

3D Research Plan

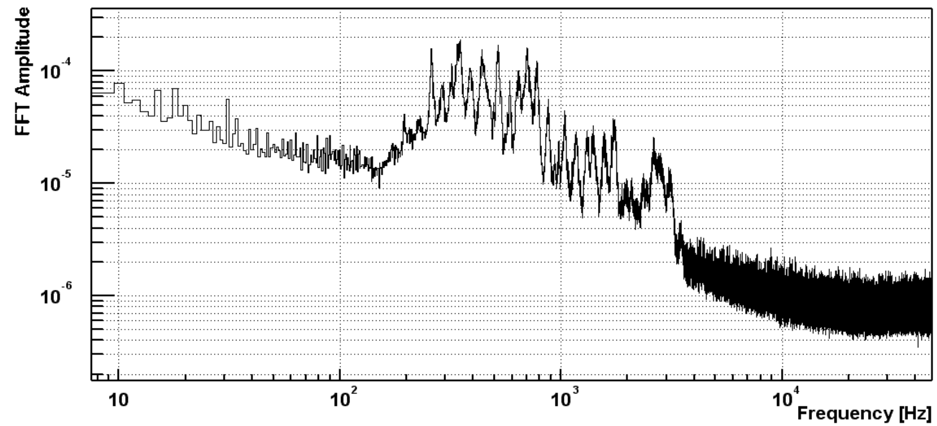
- Study data quality versus probe speed and grid spacing to optimize overall scan time.
- Study media with mould growth and other damage.
- Development of scratch correction code.
- 3D studies of disc media.
- Inteferometry study (data in-hand)

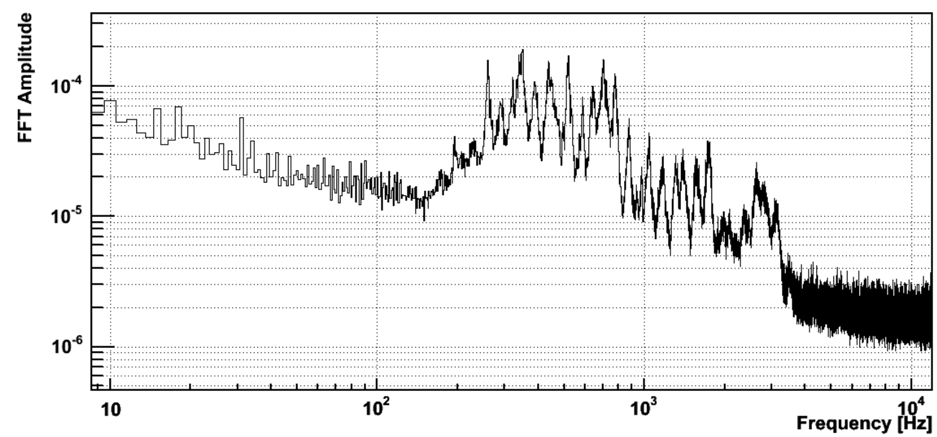
Scan Time Issues

- Factors which effect scan time
  - Probe measurements / s
    - Now at 4K /s
    - Tests have been done at 10K and 30K
    - Key issue is bright light source
  - Time sampling
    - Tests done at 96 KHz, 24 KHz, 12 KHz
  - Points across groove
    - 10 is sufficient for typical cases - how few?
- Present case
  - 4K probe + 96 KHz sample + 10 pts/groove ~ 30 hrs
- What is ultimate ("access") case?
  - 30K probe + 12 KHz sample + 5 pts/groove ~ 30 min

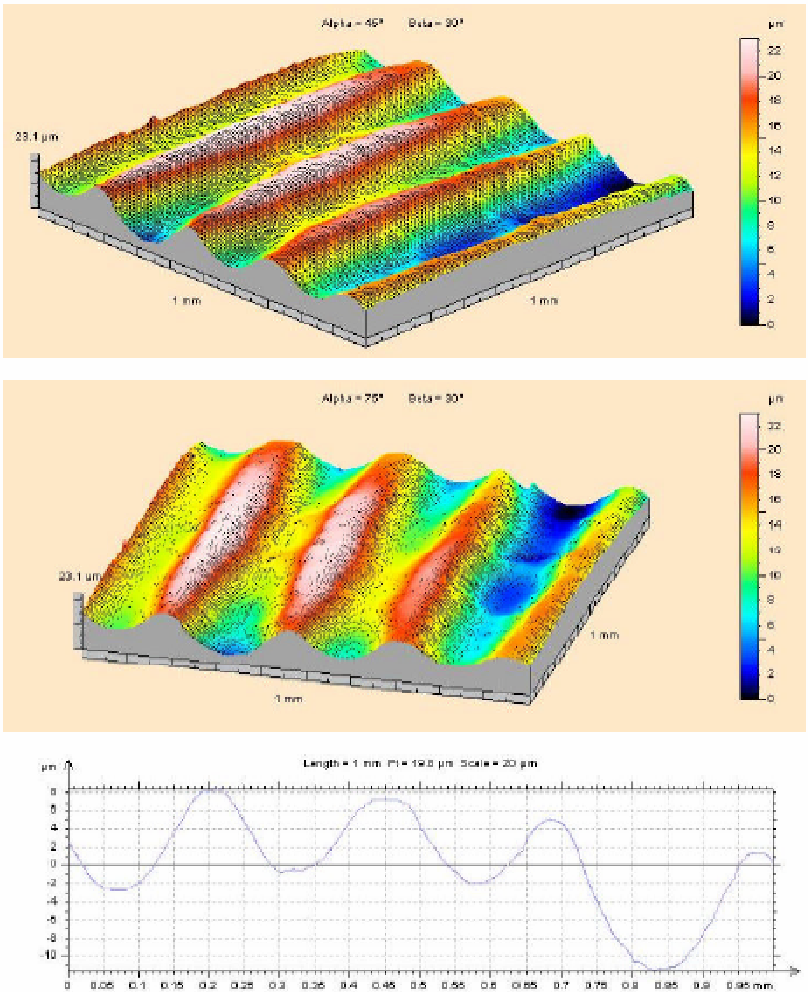
Effect of Reduced Time Sampling

- Raw data sampled at 96 KHz (5μm grid)
- Reanalyze at 24 KHz
- Increased noise in audio band due to aliasing, though not audible
- Factor of 4 gain in scan speed

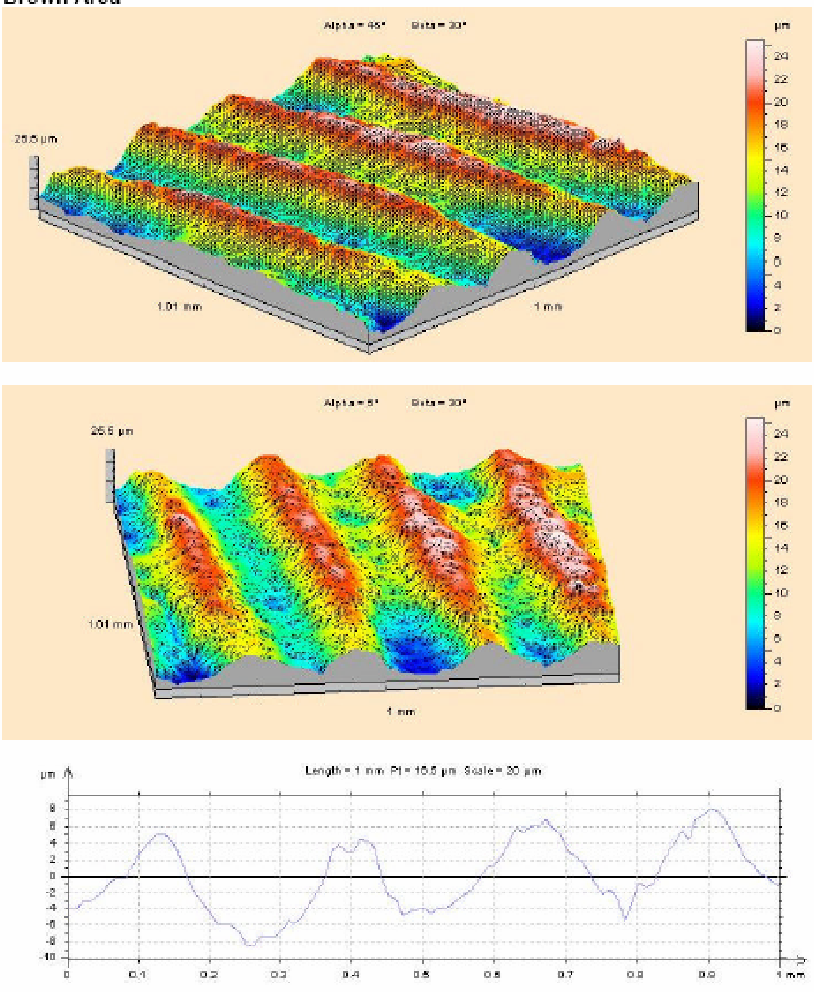




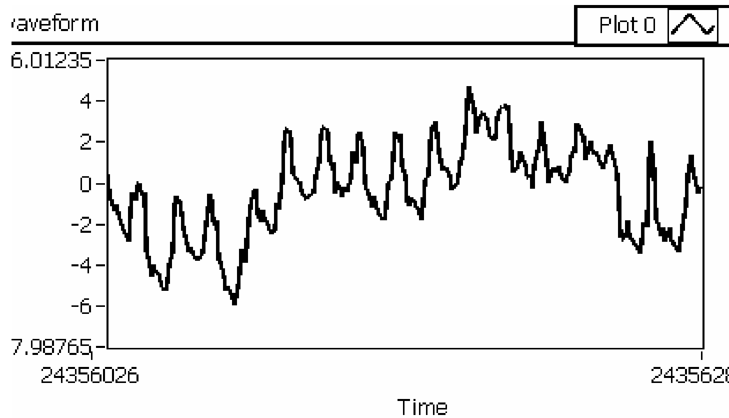
Clean Area

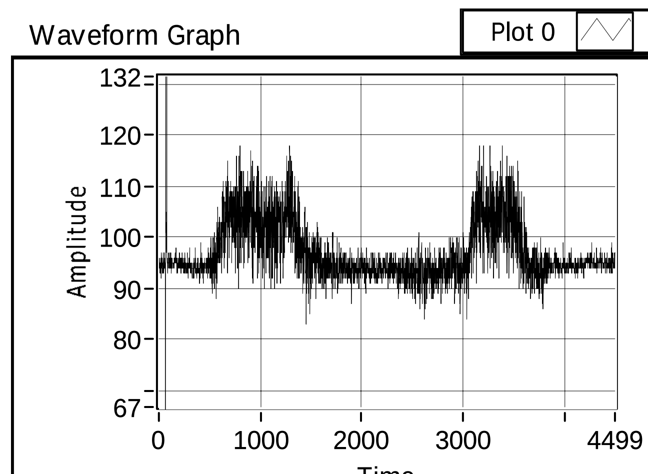


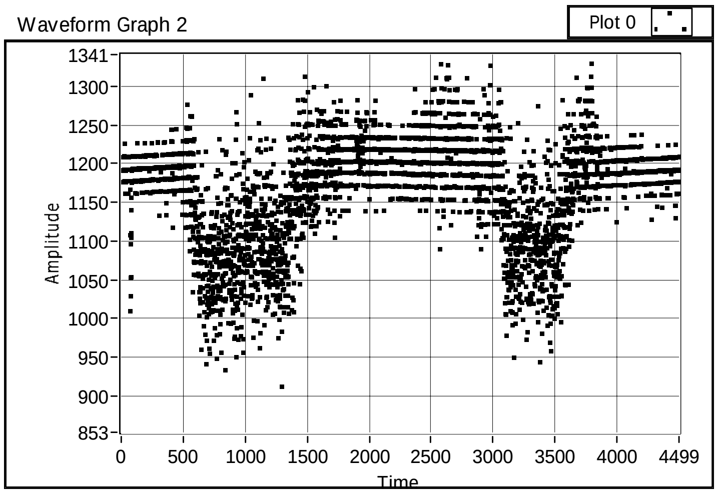
Brown Area

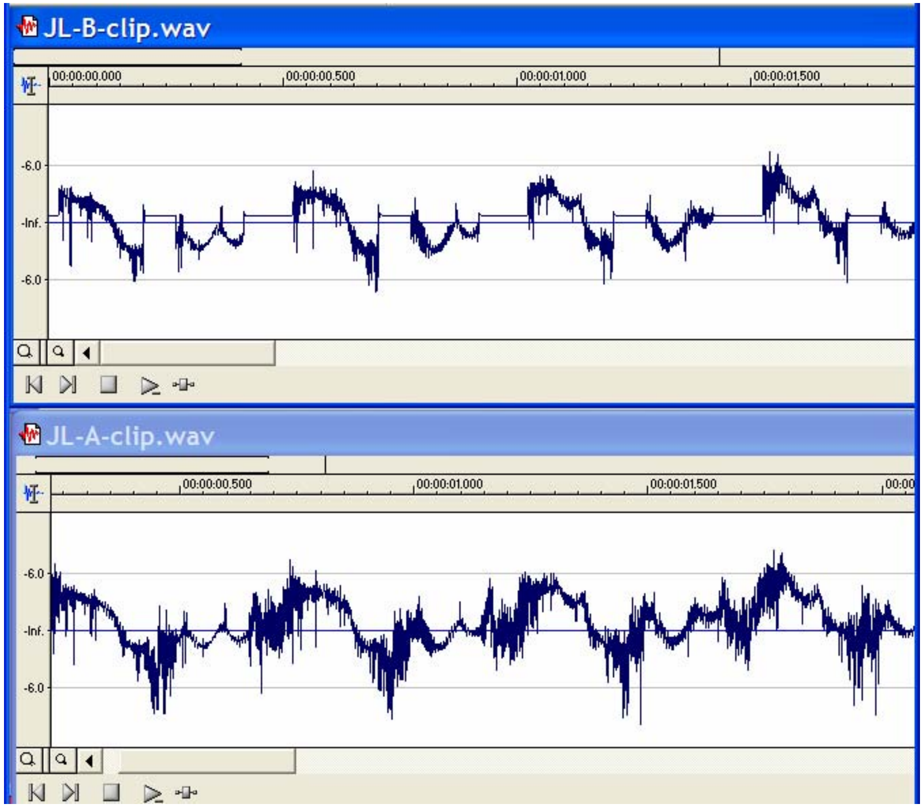


Cylinder with mould



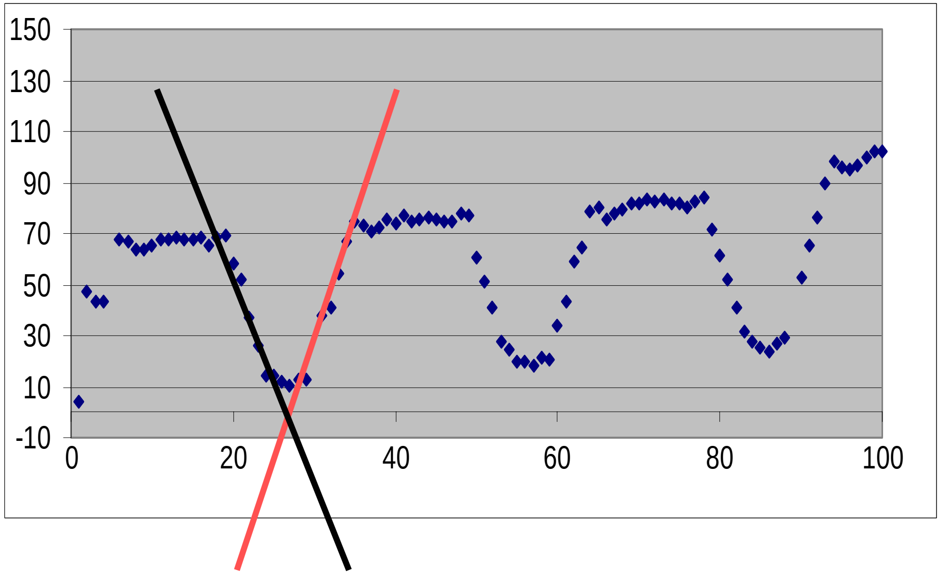



Waveform Graph 2
Plot 0
Amplitude
1341
1300
1250
1200
1150
1100
1050
1000
950
900
853
0
500
1000
1500
2000
2500
3000
3500
4000
4499
Time

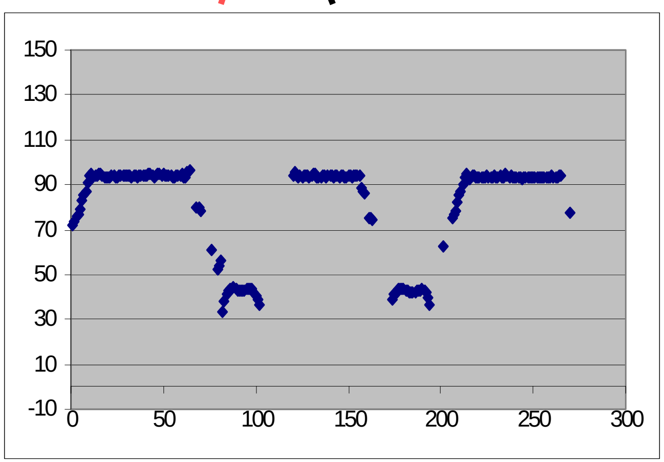

JL-B-clip.wav
00:00:00.000
00:00:00.500
00:00:01.000
00:00:01.500
-6.0
-Inf.
-6.0
JL-A-clip.wav
00:00:00.500
00:00:01.000
00:00:01.500
-6.0
-Inf.
-6.0

3D Scans on discs

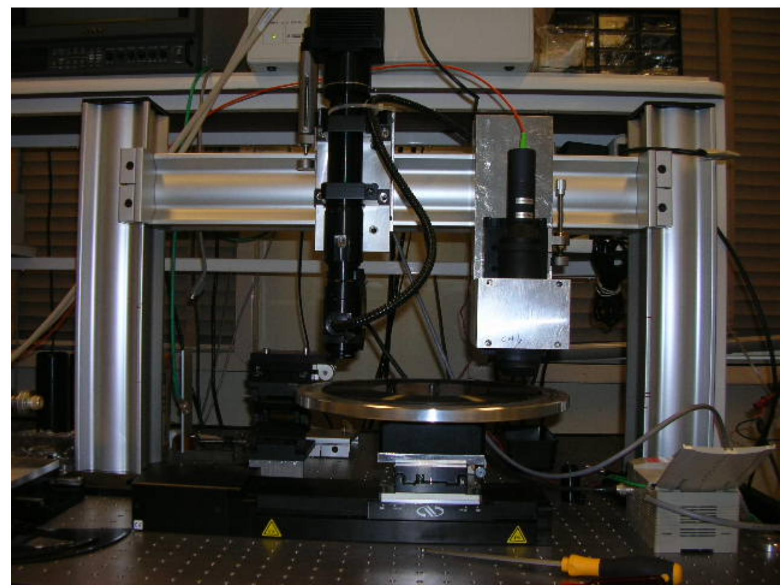
40 Hz



- Data loss on slope are key issue
- Don't expect scans as fast as cylinders
- Compare to IRENE-2D



400 Hz



Outreach and Collaboration

- The media and public have shown interest in these efforts
  - Newspaper and magazine stories (NYT, SF Chron, Trade pubs..)
  - Radio interviews (NPR, Osgood, BBC,…)
  - TV coverage, documentary inquires
  - Inquires from individuals – hundreds of emails
  - Requests pertaining to specific samples
- Invited talks ~ 25 so far
- A good vehicle to communicate the value of science, basic research, and culture. Make a point of this.

Further Collaboration and Activity

- Interest from NARA re: dictation belts
- J.McBride at Southampton UK, recv'd ~1M$ from UK gov't for research in this area, (Victoria cylinder)
  - Visit in July
  - Plans to send a PhD student to Berkeley
- O. Johnsen at EIF-Fribourg, Switzerland runs the VisualAudio project (use of films)
  - 2 students to Berkeley this summer for 3 months
- I. Fujinaga, McGill U., 3D scanning
- Invitation to visit Discoteca del Stato, Rome

From Top Quarks to the Blues
Particle Tracks Tune Up Music
Physicists Seek to Digitize Music, Restore Media
Using high-energy physics to preserve old records
Scientists find new way to play old records
Optical Metrology Reconstructs Audio Recordings
From the Higgs Boson Particle to Leadbelly
Teilchenphysik im Dienste des Kulturerbes
Teilchenphysiker retten das musikalische Erbe der Menschheit
Particle Physics Recovers Music From the Past
New technique preserves old sounds
Digitizing groovy records
De la Física a la Fonografía
Physiker retten Schellack-Aufnahmen
Particle physicists to help restore old audio recordings
How to listen to old records in the 21st century
Particle physicists rescue rare vinyl recordings
Φυσικοί βρίσκουν τρόπο να βελτιώσουν τον ήχο
Der Bosonen-Blues - Teilchenphysiker helfen alte Tonaufnahmen von Schellackplatten und Wachszylindern zu retten
Physicists find method to improve audio
Laser pour vieux vinyles
LISTENING TO RECORDS BY LOOKING AT THEM
Aus alt mach neu
Fizycy ratują stare winyle
Playing Old Records (No Needle Required)
New Hope For Old Sounds
Optical Metrology Reconstructs Audio Recordings
Digitizing the voices of the past
Science perfects sound of century-old recordings
Virtual Record Player Preserves Historic Recordings
Particle Tracking Tunes Up Music
Physicists Seek to Digitize Music, Restore Media
Groovy Pictures: Extracting sound from images of old audio recordings
How to listen to old records in the 21st century
Rescuing Recordings
REAL LIFE NEWS: PRESERVING ANCIENT RECORDINGS
Técnica permite recuperar LPs danificados pelo tempo
Inspirado na física de partículas, método digitaliza gravações sem riscos e chiados

Why I read Physics Today

Conclusions

- Image based methods have sufficient resolution to reconstruct audio data from mechanical media and reduce impulse noise.
- 2D approach may be suitable for mass digitization. IRENE will address this and other key issues.
- At present 3D methods are suitable for reconstruction of particular samples since they require ~hours per scan.
- Ongoing 3D research program addressing issues of ultimate scan time, damaged media. A 3D "IRENE" system next?
- Considerable professional and public interest, growing community
- Info at URL www-cdf.lbl.gov/~av