Phonographic Sound Extraction Using Image and Signal Processing

www.eif.ch/visualaudio

Ottar Johnsen
Sylvain Stotzer
Christoph Sudan
Stefano Cavaglieri
Frederic Bapst
Pio Pellizzari
Cédric Milan
What is the Problem?

- Records were the only mean for preserving sound until the introduction of tape in the early 50’s.

- There are huge collections of 78rpm and microgroove records around the world, including historically and culturally important sounds (radio talks, radio debates, music, …) where only a single copy exists.
What is the Problem?

- Disks, and in particular acetates and shellacs, are fragile, and for handling them special care and extensive training are required.
- Many old records are in such a bad shape that it is impossible to read them with a record player, and they are deteriorating.
- Digitized sounds must regularly be transferred to other storing medias as the existing ones become obsolete or deteriorate.
The amplitude of sound signal is stored either in the groove position for constant amplitude recording, or in the derivative of the radial groove position for constant velocity recording. Most of the disks are recorded with constant velocity, but equalized using standard curves (RIAA, NAB, FFRR, AES, …).
Fundamentals

• The shape of the groove visually represents the acoustic vibration, which corresponds to the electric signal of the recorded sound.

• The entire surface of a disk could be photographed or scanned at high resolution.
Microscope observation
The VisualAudio concept

1. A high resolution analog picture of each side of the disk is shot. The film becomes the intermediate (or final) storage media.

2. To listen to the sound, the picture is scanned using a high resolution circular scanner.

3. The sound is extracted from the digital image transforming it into a one-dimensional signal, which contains the radial displacement of the groove that corresponds to the sound.
The VisualAudio concept

1. Record  
2. Film  
3. Digitized images  
4. Sound
Why use the intermediate photographic step?

- Groove position must be estimated very accurately. But the warping of the disk exceeds the depth of field of a microscope.
- The coding of the information is the same on the disk and on the film.
- Taking a picture of a disk is a quick way to store an (almost) analog copy of the sound.
- Film is a small, cheap and a quite stable medium (more than 100 years) for storing sound information.
Grooves shapes

78 rpm groove.

33 rpm groove.
Film requirement

Requirements:

- High resolution
- Small grain
- Black and white
- Speed
Film
Shooting the picture: current system
Shooting the picture: new system

- Mettre dessin du nouveau système
Photographic illumination

- The disks are bright, their reflectivity is mainly specular: most of the reflected light has a reflective angle equal to the incidence angle.
- The best solution is to have a directional light that illuminates the disk uniformly from the lens point of view.
- A monochromatic blue light: its short wavelength improves the sharpness and fits to the spectral response of the film.
Photographic optics

- The circle of confusion $C$ is the blur caused by the depth of field ($DOF$), the focal length $f$ and the opening diameter of the lens $D$.

- The Fraunhofer diffraction produces Airy patterns: the image of a point through a lens is a spot (blur), called Airy disk.
Photographic optics

With $DOF = 1 \text{mm}$ and $\lambda = 0.41 \mu m$:

- Circle of confusion: $C = DOF \frac{D}{4f}$

- Airy disk: $d_{airy} = 2.44\lambda \frac{f}{D}$

- Total resolution: $B_{photo} = \sqrt{C^2 + d_{airy}^2} = 22.3 \mu m$
Scanner

- Glass turntable.
- 2048-sensor CCD-linear camera mounted on microscope lens above the glass.
- Light source located below the tray lightens the film by transparency.
New scanner

Mette image du nouveau scanner
Scanner

- At each rotation: scan a ring of the film.

- Adjacent rings are scanned in order to digitize the whole record.

- Sampling frequency from 25-200 ksamples/ring.

- Audio sampling frequency: 13.75-110 kHz for 33 rpm, 32.5-260 kHz for 78 rpm records.
Scanning the picture
Scanner optics

With a Numerical Aperture $NA=0.25$

- Circle of confusion: $C = DOF \frac{NA}{\sqrt{1 - NA^2}}$
- Airy disk: $d_{airy} = \frac{0.61\lambda}{NA}$
- Total resolution: $B_{scanning} = \sqrt{C^2 + d_{airy}^2} = 13 \ \mu m$
Processing the image

- **Principle:** Estimating the radial position of the groove

- **How:** By detecting the edges of the grooves (2 or 4) if possible with subpixel accuracy

- **Problems:** Noise and distortions
Processing the image

Noise and distortions:

- Dust and defects on the record
- Non constant illumination
- Film grain
- CCD noise
- Sampling time jitter
- Scanner vibrations
- Bad centering
Processing the image
Processing the image
Processing the image
Processing the image
Processing the image
Extracting the sound
How to measure the quality?

- Perceptual
- Measure noise in silent sections of a groove
- Measure the SNR with a sinewave from a test record
- Analyzing the spectrum can help finding the distortion or noise causes
Signal to noise analysis
Signal to noise analysis

What is the resolution needed to reach a 40 dB SNR on the extracted sound of a 78 rpm?

Hypothesis:

- Sound bandwidth
  \[ B_s = 12 \text{ kHz} \]
- Noise bandwidth
  \[ B_n = 130 \text{ kHz} \]
- Maximal groove deviation
  \[ A = 75 \text{ \mu m} \]
- 78 rpm disk => 4 groove edges
Signal to noise analysis

Resolution needed to reach a 40 dB signal to noise ratio:

- Constant amplitude records: $1.75 \mu m$
- Constant velocity: $0.25 \mu m$
- Equalized records (RIAA): $1.28 \mu m$
Signal to noise analysis

• Does it make sense:
  A blur of 30 μm, and an resolution of 1 μm?

• Yes! A blur is a low pass filtering
  The resolution is related to noise!

• But decreasing the blur helps
Advantages

• Sound retrieval without contact.
• Fast extraction time.
• Disks in virtually all conditions (even delaminated, broken, deformed, etc.) can be read and the sound restored.
• Each and every disk format (size, speed, cutting, etc.) is read using the same equipment.
• Image processing is something very well established. It is relatively easy to make all kind of corrections to the physical incoherencies of the disk.
Similar work

Dr. Haber et al. at Lawrence Berkeley Labs are working on a system to directly scan the records and cylinders using a microscope visual inspection system.

- They need several hours for scanning.
- In principle, they should get better result in particular with their 3D system.
- Their system can be used to save selected records, not for a mass saving.
Conclusions

- The concept of VisualAudio has been demonstrated: it is possible to extract the sound from a record by putting it on film, scanning it and processing it.
- This project is not a final product. The quality is not yet satisfactory, but the path to improvement is known and new prototypes are currently been built.
- Improvements are needed and underways in several area: photography, mechanics, optics, signal processing, pattern recognition,…
Conclusions

- New prototypes are currently been built
- In one year, we plan to be able to start large scale record saving
Collaboration

- École d’ingénieurs et d’architectes de Fribourg
- Swiss National Sound Archives
- University of Fribourg
- École d’arts appliqués de Vevey

With funding from:
- Gebert Ruf Foundation
- Swiss National Science Foundation
More informations

You can listen to the sound, see a video demonstration, and get more information at:

www.eif.ch/visualaudio