Optical Retrieval and Storage of Analog Sound Recordings

Stefano S. Cavaglieri – Prof. Ottar Johnsen, Prof. Frédéric Bapst
Fonoteca Nazionale Svizzera - Ecole d'Ingénieurs de Fribourg
Lugano 6906, Switzerland - Fribourg 1705, Switzerland

The optical retrieval and storage technique described in this paper provides a way for transferring sound information from an analog disk to a visual carrier avoiding any mechanical contact. The process is straightforward: we take a picture of each side of the disk using a dedicated analog camera, we store the film as our working copy and when needed, we scan the film and process the image in order to extract the sound.

FUNDAMENTALS

Sound is part of the natural environment of mankind. It is first of all a perceived phenomenon provided by the sense of hearing. It allows spoken communication between people, but it is also music, which is another form of expression and communication, and noise, a byproduct of today's industrial society.

Sound is analog by nature. In fact, an acoustic vibration is a movement of particles around an equilibrium position in an elastic medium.

The ability to communicate is a privileged bond between people; thus we learned how to capture, to retrieve, to transmit, to store, and to reproduce sound.

How do we capture and reproduce sound?

We capture and reproduce sound by using several sorts of transducers. This is best explained taking a microgroove disk as our real world example:

Recording

The singer goes to a recording studio and sings into a microphone. The microphone transforms the acoustic vibration into an electrical signal. The signal flows through a number of electronic devices reaching a tape recorder. The tape recorder transforms the electrical signal into a magnetic field. The magnetic field is applied to a magnetic tape.

Cutting

The magnetic tape is read by a tape recorder. The tape recorder transforms the magnetic field into an electrical signal. The signal flows through a number of electronic devices reaching a disk recorder. The disk recorder transforms the electrical signal into a mechanical displacement of the cutting stylus. The cutting stylus cuts the groove into the surface of the lacquer.

Reproducing

The reproducing stylus reads the groove of the disk (the equivalent of the lacquer). The mechanical displacement of the stylus is transformed into an electrical signal. The signal flows through a number of electronic devices reaching a loudspeaker. The loudspeaker transforms the electrical signal into an acoustic vibration.

Of course sound can be digitized and/or manipulated, but this will only add or modify part of the above chain, while the source (acoustic vibration → electrical signal) and the target (electrical signal → acoustic vibration) remain the same.

What do we expect when reproducing recorded sound?

The most immediate desire everyone has, is being able to reproduce the same, or at least the closest readability, quality, ambience, emotions, etc. as when that sound was originally generated.

FACTS

Cutting a disk was in practice the only mean for preserving sound until the introduction of magnetic tape in the early 50's.

There are huge collections of 78 rpm as well as microgroove disks around the world, and there is an increasing demand of republishing those materials.

Disks, and in particular acetates and shellacs, are fragile, special care and extensive training are required for handling them.

As of today, only few manufacturers still produce turntables, heads, styluses, and replacing parts.

The majority of all archives and collectors have no time, no personnel, and no money to keep their collections in shape. The main reasons being that copying can only be done in real time and that no "modern" sound carrier is designed for long term storage.

Digitizing a sound using standard specifications (44.1kHz sampling frequency, etc.) does neither improve the quality nor does it extend its life. On the contrary, the continuous changes and additions of new standards potentially reduce the life cycle of a particular format.

The reproduction of a stored digital signal is much more complex than the reproduction of an analog signal.

... a little bit scary, isn't it?

GOALS

Let's look at all those issues one at the time, and set some goals. There are several topics that could be raised, but the main target will focus on:

- retrieving the information from an analog disk as an "image" without touching its surface;
- processing the image captured off the disk surface and recognize its content electronically, transforming it into an electrical signal;
- reducing the size of the image by a high factor, and store it on a stable carrier for long-term archival purposes;
- reducing the overall processing time significantly (compared to the real playing time);
- reducing the overall cost significantly.

THE BASIC IDEA

Disks were first monophonic; the groove was cut either vertically or laterally. After a while disks became stereophonic, the size of the groove was reduced and cut 45° / 45°.

The shape of the groove visually represents the acoustical vibration, which corresponds to the electrical signal of the recorded sound.

The entire surface of a disk could be photographed or scanned at high resolution, then reduced using a technique similar as for printing integrated circuits.
The resulting micro-image could be burned as a hologram on something like, as an example, a small crystal.

A laser beam could finally read the image stored on the new carrier.

THE VISUALAUDIO CONCEPT

The initial concept, for prototyping the above idea, is based on the fact that when observing a sound recording on disk, using a microscope, we can "see" the sound. In fact, the groove that contains the sound is quite visible. If it is visible, it means that the sound information is contained in the image of the record. It is the radial displacement of the groove that contains the sound. This concept lead to 3 steps (which are shown on Fig. 1) to work on:

1. An analog picture of each side of a disk (either 33 1/3 or 78rpm) is shot. The film must have a high spatial resolution and be relatively large, since we wish to catch the finest details of the groove. In fact, the film must therefore be of about the same size as the record. This process can be done quickly. The film is cheap, and can be stored for a long time (more than 100 years). That way, the sound information is preserved in case the original disks deteriorate.
2. When anyone wants to recover (to listen to) the sound, the film, with the picture of the disk, is scanned using a specially designed rotating scanner, and digitized. At this point a digital image of the record is stored.
3. The sound must then be extracted from the image. This requires image-processing techniques in order to extract the radial displacement of the groove (which contains the sound), to detect cuts and to correct other defects. Digital signal processing must be applied to the recovered sound in order to improve the sound quality.

The above concept was verified thanks to several student projects [1], [2] and [3].

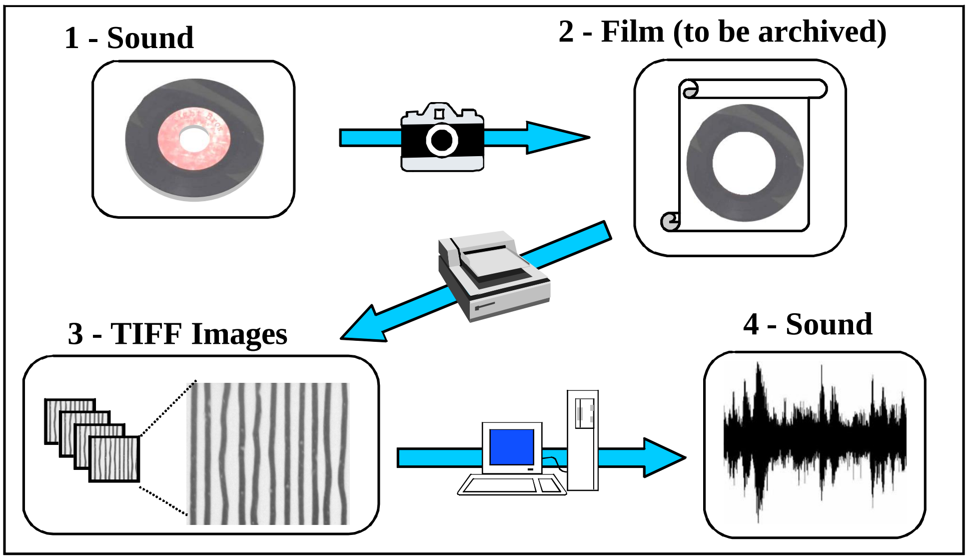

Fig. 1: the VisualAudio concept

DISK CHARACTERISTICS

The typical disk characteristics, which are useful for this project are:

78 rpm records

- Groove width: 31-187 µm
- Groove deviation: 75 µm
- Distance between grooves: 265 µm
- Bandwidth: 100-12000Hz
- Signal to noise ratio: 32-40 dB
- The bottom of the groove is flat

33 1/3 rpm records

- Groove width: 25.4 µm
- Groove deviation: 28 µm
- Distance between grooves: 85-125 µm
- Bandwidth: 30-16000Hz
- Signal to noise ratio: 45-60 dB
- The groove has a triangular shape

PICTURE TAKING

With such tiny groove deviations, 75 µm (78 rpm) or 28 µm (33 1/3 rpm), particular care is a must in the picture taking phase. High resolution films have a resolution of about 600 lines, corresponding to 1200 dots/mm. Taking into account the optical degradations, the resolution is about 1.3 µm. With such a resolution, the optics and the conditions for shooting the picture must be chosen carefully. The film must have about the same size as the record. Our pictures were made in a photo lab.

PICTURE SCANNING

The ideal scanner would be a rotating scanner. In a first setup, we used a digital video camera as scanner. The disadvantage is that we are getting geometric distortions. In a new setup we are using a linear scanner combined with a rotation of the film. Figures 2 and 3 show the picture obtained from 33 1/3 and 78 rpm disks. We observe that for the 33 1/3 rpm record, we see a single groove, while for the 78 rpm record, we see 2 parallel lines corresponding to the 2 sides of the groove, while the flat bottom is black. Note that the brightness should be the opposite, but we are working with negatives. It is clear that image processing is necessary to recover the sound

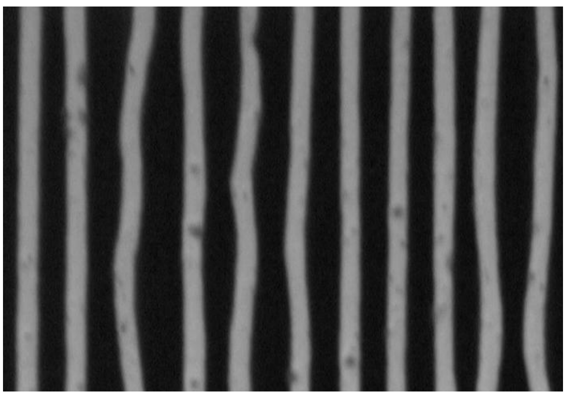

Fig. 2: picture of a section of a 33 1/3 rpm record

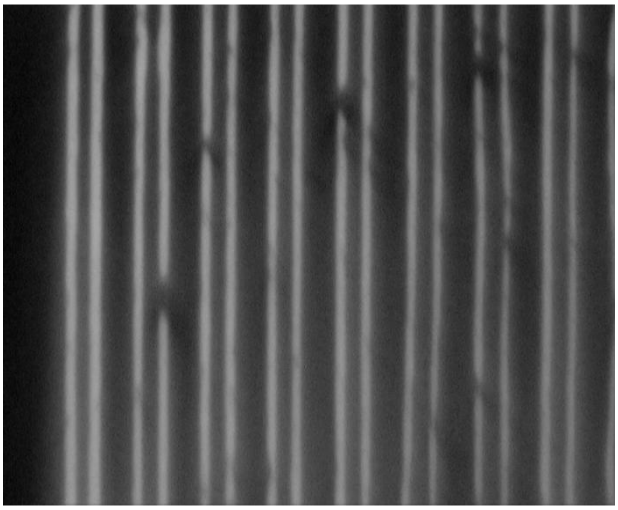

Fig. 3: picture of a section of a 78 rpm record

IMAGE PROCESSING

The image must be processed in order to extract the sound from the scanned picture. The main steps are:

- Extract the groove borders with (if possible) a subpixel accuracy
- Suppress spurious points
- Avoid cuts in the groove or a merge of neighboring grooves

Figure 4, 5, and 6 illustrate the image processing step

Fig. 4: binarized picture with a fixed threshold

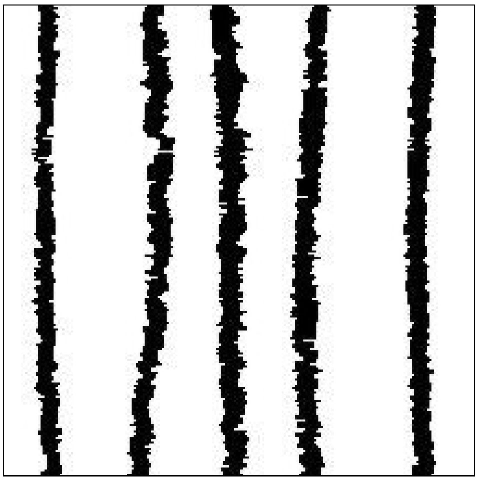

Fig. 5: binarized picture after adaptive threshold

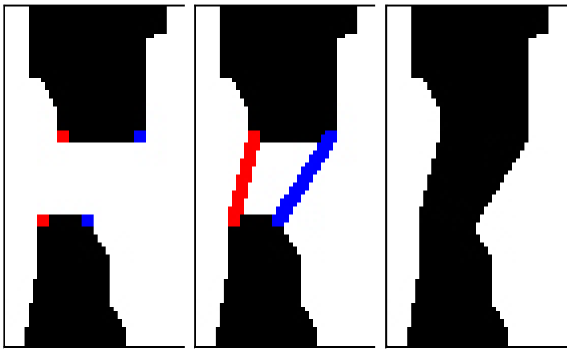

Fig. 6: cut suppression technique

Due to irregularities in the groove illumination, the threshold to detect the groove position should be adaptive, if not, additional cuts and merges of grooves will appear. Spurious points might appear when processing the pictures. Such points can be easily cleaned. Cuts in the groove can be corrected using interpolation. Merges of grooves can be processed the same way. With 78 rpm records, to have 2 white lines for each groove can be helpful.

EXTRACTING THE SOUND FROM THE GROOVE POSITION

The sound signal is obtained by extracting the radial position of the groove. Often, the sound is oversampled (the sampling frequency depends on the tangential distance between pixels). In an example, the sound was sampled at 197 kHz. As the sampling rate necessary is only 30 kHz for 33 1/3 rpm (15 kHz bandwidth), There is an oversampling by a factor of 6. This allows a reduction in the noise since most of it is above 15 kHz.

SIGNAL TO NOISE CONSIDERATIONS

Getting a good signal to noise ratio is the most important part in this project. There are many sources of noise. The most important are:

- The record might be dusty, or the groove border is dented.
- The optics for the photography and the scanning can lead to distortions.
- The film (the granularity) causes noise.
- The scanning process, means that the radial position of the groove is quantized.

The scanning process must also have a high resolution Here also, a resolution of about 1µm is necessary. With such resolution we are near the light wavelength, and monochromatic light must be used.

RESULTS

An extract of music will be played. We would estimate that the signal to noise ratio is about 20 dB. The quality must be improved, but the concept has been demonstrated: it is possible to extract the sound from a record by putting it on film, scanning and processing it.

The processing time was at the beginning a limiting factor. Extraction of a few seconds of music took several hours. By replacing the Matlab program by a C++ version, a speedup by about a factor 60 was obtained. Further improvements are possible.

This short project is not a final product. The quality is not yet satisfactory, but the path to improvement is known.

Much work must still be done, in particular concerning the understand of the properties of shellac and vinyl recording, the picture taking, the scanning, the groove extraction, the image to sound conversion.

ADVANTAGES

- The picture is shot without interfering with the surface of the disk. There is no need to manipulate the disk except for placing it on the photo stand.

- Disks in virtually all conditions (even delaminated, broken, deformed, etc.) can be read and the sound can be restored.

- Each and every disk format (size, speed, cutting, etc.) is read using the same equipment.

- Image processing is something very well established so far. It is relatively easy to make all kind of correction to the physical incoherencies of the disk.

- Film (as an intermediate step in the development of the basic idea) is a quite stable, small, and cheap carrier for storing sound information. This means that it might be implemented as well as a long term storage format.

DISADVANTAGES

There are not too many industries interested in developing a niche product.

REFERENCES

[1] Glatz, S., Milanovic, S. "VisualAudio", Semester project at Ecole d'ingénieurs et d'architectes de Fribourg, Switzerland, 2000.

[2] Chassot, O., Miauton, C.,"VisualAudio", Diploma project at Ecole d'ingénieurs et d'architectes de Fribourg, Switzerland, 2000.

[3] Bosi, N. O.,"VisualAudio: acquisition de l'image", Diploma project at Ecole d'ingénieurs et d'architectes de Fribourg, Switzerland, 2000.

THANKS

The authors would like to thank Thierry Fumey for taking the pictures of the records, Laurenz Altwegg for his advice, Olivier Chassot, Nicolas Bosi, Cédric Miauton, Sylvain Glatz and Sladjan Milanovic for their help in this project.

This project was partly funded by project HES-SO-01.01TI

Address of the authors:

S. Cavaglieri, Fonoteca Nazionale Svizzera, Lugano, Switzerland

cavaglieri@fonoteca.ch

Ottar Johnsen, Ecole d’ingénieurs et d'architectes de Fribourg, Switzerland

johnsen@eif.ch

Frédéric Bapst, Ecole d’ingénieurs et d'architectes de Fribourg, Switzerland

frederic.bapst@eif.ch