DETECTION OF THE GROOVE POSITION IN PHONOGRAPHIC IMAGES

Ottar Johnsen, Sylvain Stotzer, Frédéric Bapst – Rolf Ingold

College of Engineering and Architecture of Fribourg – University of Fribourg, Switzerland
ottar.johnsen@hefr.ch, sylvain.stotzer@ch.delarue.com, frederic.bapst@hefr.ch, rolf.ingold@unifr.ch

ABSTRACT

The sound of phonographic records is contained in the radial position of the groove (and sometimes also its depth). The sound is usually extracted by a mechanical device (the stylus of the record player). Many old records are in such bad shape that no mechanical contact is possible. An alternative is to take a picture of the record and to estimate the radial groove position, using image processing. In this paper, we present the technique to extract the radial position of the groove and therefore the sound of the records.

Index Terms— *Image processing, image edge analysis, pattern recognition*

1. INTRODUCTION

On record players, a needle follows the groove's radial position and converts it into an electrical signal corresponding to the sound. Therefore, the image of the record contains the sound. This observation has lead to several proposals of systems to extract the sound from records using image processing [1]. Our proposal, called VisualAudio is a three-step concept [2][3][4][5]:

1. An analog picture of a disk is taken, in order to preserve the sound information in case the original record deteriorates.
2. When one wants to recover the sound, the film is digitized using a specially designed rotating scanner.
3. The sound is then extracted from the digital image by measuring the radial displacement of the groove.

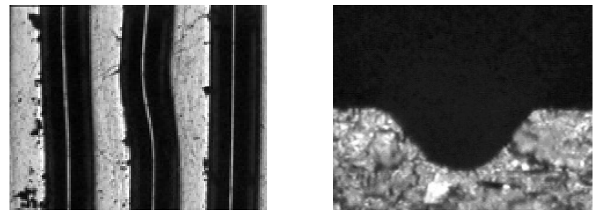
While our previous publications described the general concept and the signal to noise analysis of the VisualAudio system, this paper will focus on the third step of the process: the image processing algorithms used to extract from the scanned image the groove position and therefore the sound.

Step 2 may be skipped, as one might directly scan the record. However, archivists like to keep an analog copy. Furthermore, as time is a critical issue in such an archiving system, the challenge is to save a large amount of records quickly, before their complete physical destruction. Taking a picture of the disks is a quick way to store an (almost) analog copy of the sound content in its current stage of conservation. The sound extraction could then be done on demand.

2. IMAGE ANALYSIS

2.1. Disk characteristics

The amplitude of the sound signal is stored in the groove position, represented by the radial displacement of the stylus on a turntable. Figure 1 shows a picture of a small section of a record as seen on a microscope and the shape of the groove.



(a) (b)

Fig. 1. Top view (a) and profile (b) of 78 rpm records.

	78 rpm	**33 rpm**
Groove width:	31-187 µm	25.4 µm
Groove deviation	75 µm	28 µm
Groove spacing	200-300 µm	85-125 µm
Bandwidth	100-12000 Hz	30-16000 Hz
Groove shape	round	triangular

Table 1: record characteristics [6].

Table 1 shows some record characteristics that influence the image processing. The groove's shape depends mainly on the kind of disk. 78 rpm (rotations per minute) records have round grooves. Microgroove 33 rpm records have triangular grooves; combining the displacement of both walls can be used to store 2 channels for stereophonic records. The groove's shape has an impact on the disk picture. When illuminated and shot perpendicular to the record plane, triangular grooves don't reflect light and therefore appear as a unique trace on the negative film. With the round grooves of 78 rpm records, the walls don't reflect light, the bottoms of the grooves do, leading to two traces.

2.2. Image characteristics

During the scanning, the film lies on a glass turntable. A 2048-sensor CCD-linear camera mounted on microscope optics is fixed above the glass and a light source located below the tray lightens the film by transparency. During each rotation of the glass turntable, we scan a ring of the film, whose width depends on the optics magnification. By radially displacing the tray, adjacent rings are scanned in order to digitize the whole record. The sampling frequency is 65.5 or 131 ksamples per ring. This rate combined with the disk rotation speed defines the audio signal sampling frequency of 36.4 or 72.8 kHz for 33 rpm records, and of 85.2 or 170.4 kHz for 78 rpm records. The circular scanner has the advantage of transforming the circular disk picture into a rectangular image, thus avoiding a coordinate transformation. Therefore, we scan in polar coordinate. The image of the whole record is a matrix, where one axis corresponds to the radial position, and the other axis to the angular position. Considering the record, one axis corresponds to the groove deviation, the other to the time. Figure 2 shows a section of the digitized image from a 78 rpm record, where the groove is depicted by two traces.

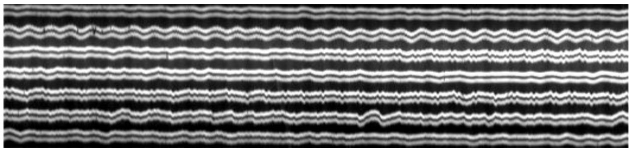


Fig. 2. Section of a ring image acquired on a 78 rpm record.

3. GROOVE EXTRACTION

The groove extraction algorithm must accurately estimate the radial groove position since that position corresponds to the instantaneous sound. For the algorithm description, we consider the 78 rpm records with 2 white traces, and therefore 4 edges, but it is straightforward to extend it to the case with a single trace and only 2 edges.

As seen on Figure 2, the groove is almost parallel to the tangential direction (except for the sound modulations). Therefore, if we make a vertical cut in the picture, we get the radial intensity profile as shown on Figure 3 (top part) for one of the grooves.

We observe in the intensity profile 2 high intensity regions, corresponding to the 2 white traces. The groove radial position is determined by the 4 edges. The algorithm will extract the 4 edges and combine them to estimate the instantaneous groove position and therefore the sound. The consecutive groove positions will be combined to get the total sound. In order to get the most accurate estimation of the edge position, a priori knowledge of the groove picture must be considered:

- The groove has physical properties as outlined in Table 1.
- The groove's width is constant for mono records, meaning that the displacements of all of the groove edges are equal.
- The sound properties define boundaries for the groove displacement between consecutive samples.

Due to the degradations, the edges are represented by blurred areas that can be described by four parameters [7]:

- Location
- Transition width
- Dark intensity I_d
- Light intensity I_l

The groove or sound extraction algorithm can be divided into 3 parts:

a) Coarse edge position estimation
b) Fine edge position estimation
c) Groove position estimation and the sound reconstruction including the suppression of artifacts.

Phase a) and b) use one-dimensional edge detection. It is only in phase c) that we take into account the two-dimensional groove properties. Notice that in order to get a good groove position, and therefore sound quality, subpixel accuracy is required.

3.1. Coarse edge position estimation

The coarse groove position estimation is necessary in order to avoid making wrong decisions, in particular for noisy images, where there might be several false edges due to noise or dust. For that reason, only wide enough edges will be detected. This is obtained by convolving the image with a double box filter *b(x)* having a $\lambda \times 1$ kernel:

$$b = \left[-1 \ldots -1\ 0\ 1 \ldots 1\right]$$

where λ is usually between 10 and 20. The result is a smooth approximation of the derivative of the intensity profile, as shown on Figure 3. The extrema are the coarse approximations of the edge positions.

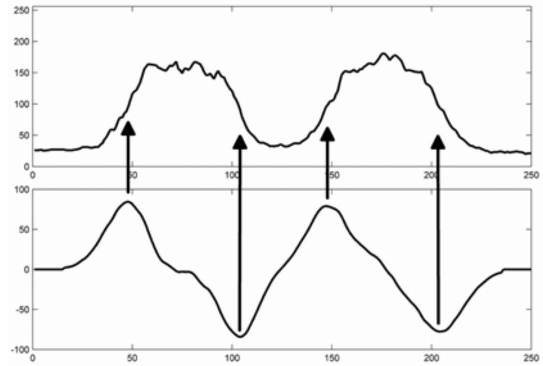


Fig. 3. Intensity profile of a 78 rpm record (top) and smooth derivation of the intensity profile (bottom).

3.2. Fine edge detection

We must determine the first edge parameter (location) by using the other three, which could vary in time due to local conditions. As long as the blur is homogeneous (not saturated) in the radial direction for each sample line, the detection algorithm has to detect the middle of the blurred area. This could be determined using the extrema of the first derivative of the intensity profile, or from the zero-crossings of the second derivative. But derivatives are sensitive to noise; therefore we propose another method to detect the middle of the transition.

The derivative of the intensity profile is approximated by convolving the intensity profile with a box filter. The size of this box must be large enough to filter the noise and guarantee the uniqueness of each solution, but it is limited by two factors: the size of the transition and the intergroove (space between grooves) spacing. This derivative is used to define the regions of interest for each groove, and the most relevant points for the transition fitting.

The ideal groove intensity profile can be modeled as two boxes for a 78 rpm record (or one box for a 33 rpm record). The blurred groove can then be modeled by trapezoids, where the top and bottom are defined respectively by the lowest grey level I_d of the space between the traces and the highest grey level I_l of the current trace. Sides are approximated by least squares fitting of the points having the higher dynamic, based on the derivative (Figure 4). The edge is then located at the middle of the segment delimited by the crossings between the sides and I_d and I_l. Interpolation is used to get subpixel accuracy. Notice that edges comes in pairs: raising and falling edges.

3.3. Groove position estimation

The traces are obtained by looking among the edges detected in the previous step, the best candidates for the raising and falling edges. Several cases appear: for some time samples, there are no edge point candidates, while in other cases there are several candidates. When there are no candidates, the edge point is predicted using the other edges of the groove. When there are several candidates, the best is selected by minimizing the distance to the estimated position, as shown on Figure 5.

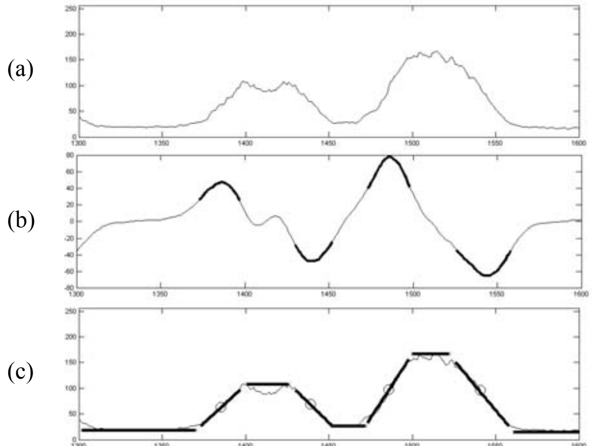


Fig. 4. (a) Intensity profile from a 78 rpm groove. (b) First derivative of the profile. (c) Groove modeling by least squares fitting of the bold parts from the derivative. Edges are indicated by the circles at the middle of the transitions.

The traces must be combined to detect the groove position as a function of time. Usually, the groove position is the average of the 4 edges of the groove. In the case of artifacts in the grooves, special algorithms are used.

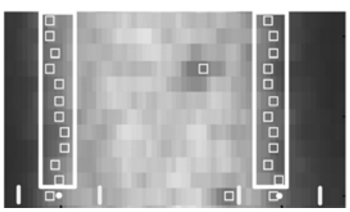


Fig. 5. Example of selection of the edges points.

3.4. Local damage restoration

Record degradations introduce noise, in particular impulsive noise. A small part of a damaged record is shown on Figure 6. There are many audio restoration techniques that reduce such noise [8], but the image allows often a more intelligent restoration compared to using only the 1-D sound.

The corrupted map system is used to locate and correct the effect of record damages [5]. The basic assumption is that the groove image degradations produce replacement noise: a dust or a scratch will replace the value of the corresponding pixels on the digital image without any correlation to the recorded sound.

The detection step consists in locating these spurious patterns, and building a map of the corrupted pixels they produced on the digitized image (Figure 7). As the traces are almost parallel to the tangential axis of the image with very smooth radial moves, the detection step is performed using a tangential gradient to locate the spurious patterns. Due to the non-isometric format of the image, these spurious patterns appear mainly as elongated spots in the radial direction on the acquired rings and are then well located by the tangential gradient. Due to the blur, the pixels surrounding the patterns are also corrupted by additive noise: blur produces a smoothed area where pixel values are a combination of the real pixel values and of the corrupted area values.

Once the map is complete, we must try to correct these corrupted pixels or avoid using them during the image processing step.

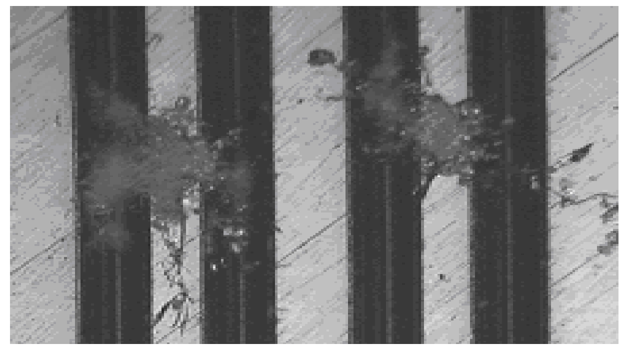


Fig. 6. Microscope view of a record damaged by fungus.

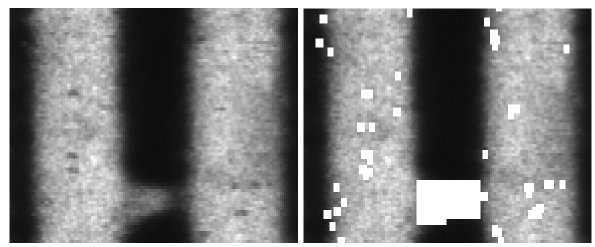


Fig. 7. Portion of a groove and version with corrupted pixels marked.

3.5. Restoration of highly damaged records

In case of large degraded areas with smooth edges, it is hard to accurately localize the damaged areas on the image with the corrupted pixel map. Any edge detected in these areas would be inaccurate. If the scratch's width and radial displacement over time are similar to the ones of a trace, it may also produce false edge detections and lead to a trace cut or a merge with another trace.

A way to separate the traces is to use only the low frequency components of the groove displacement. These low frequencies are driven mainly by the spiral of the record and the off-axis of the film on the scanner. The proposed solution is a two passes trace extraction algorithm [5]. The first pass edge detection is applied on images which have been strongly averaged in the tangential (time) axis. On these smoothed images, the traces look to be almost tangential with low frequencies modulations, as shown on Figure 8. The scratches and other large spots on the image have been smoothed and will not mislead the trace following algorithm anymore. The true trace edges are located in a close distance from the smoothed edge, bounded by the maximum amplitude of the sound frequencies which have been removed on the smoothed image.

The second pass edge detection is then applied on the original acquired image, but the ranges used for the trace following are then defined by the first pass trace extraction results, and only the edges points within the range will be considered.

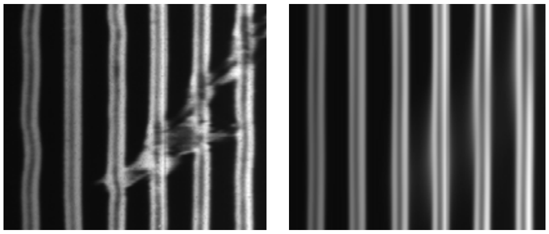


Fig. 8. Degraded portion of a groove and it's smoothed version.

3.6. Groove to sound conversion

Several corrections, which are described in more details below, must be performed to recover the original recorded sound.

In records, the sound is not contained in the radial groove position, but in the radial velocity. It means that the radial position is the integral of the sound. Therefore, to recover the sound, we must take the derivative of the groove position.

The radial groove position contains 2 components that aren't part of the sound: the ramp due to the spiral of the groove, and a periodic (nearly sine wave) component with a frequency of 78 or 33 rpm due to the bad centering of the records. A high pass filter with a 20 Hz cutoff frequency suppresses both components.

To prevent the groove excursion for low frequency sound from being too large, and to the high frequency components from being buried in noise, equalization was applied on the sound signal before recording which attenuated low frequency and boost high frequency. The inverse must now be applied to restore the sound, similarly to a turntable pre-amplifier [6] [9].

Records, in particular the oldest, have a limited bandwidth. A low pass filter can be applied to suppress out of band noise.

All these corrections are combined into a single digital filtering of the groove position to get the sound.

Finally the sampling rate is converted to the standardized 44.1 or 48 kHz frequency.

4. RESULTS

Various records have been used to test the performances and to tune the algorithms. They include both regular records and test records containing silent and sine waves groove sections. Except for records in very bad shape, the groove position and the sound were successfully extracted. The main results are.

- The signal to noise ratio obtained for good quality 78 rpm records is about 19 dB using the standardized NAB conditions [6]. Often the sound level is higher than the standard and the signal to noise ratio would then be better.
- The total harmonic distortion (THD) is about -44 dB.
- For good quality records, using regular turntable is better.
- For records in a shape requiring a contact less reading, this solution is very interesting. The sounds of several previously unreadable records have been extracted.
- The standard deviation of the edge position error is 0.4 µm corresponding to 0.2 pixels.
- Increasing the image size by using in the scanning 131 at place of 65.5 ksamples per rotation or using a 10x magnification at place of 5x doesn't improve the quality.
- The quality varies strongly from a record to another and sometimes from one part of a record to the other.
- On some records, only the outer edges are of good quality, on others, only the inner.
- Complex edge detection (and signal restoration) techniques haven't led to significant improvements compared to those presented in this paper.
- The processing time is roughly equal to the sound duration in spite of no optimization effort.

5. CONCLUSION

An algorithm to detect accurately the groove position from the image of a record has been presented. It makes the contact-less extraction of the sound from phonographic records possible. It is especially useful for disks in bad condition, for example delaminated or broken records, where no techniques allow today its extraction. We hope that this technique can help to preserve the sound archives, which are part of our cultural heritage. Several recovered sound samples can be found at www.eif.ch/visualaudio.

The image processing technique presents numerous advantages over the groove following methods (turntable, optical fibre, laser reader ...), as many groove degradations can already be detected and corrected at the image level, therefore lowering their impact on the extracted sound. We are continuously improving the algorithm for records in very bad shapes, in particular for records with shrinkage of the lacquer coating.

ACKNOWLEDGEMENTS

The authors thank T. Fumey and C. Milan for their contribution. This project is funded by the Gebert Rüf Foundation, project HES-SO-01.01TI, and the Swiss National Science Foundation.

References

[1] V. Fadeyev, C. Haber, "Reconstruction of Mechanically recorded Sound by Image Processing", Journal AES, Vol. 52, No. 12, pp 1172-1185, 2003.

[2] S. Cavaglieri, O. Johnsen, F. Bapst, "Optical Retrieval and Storage of Analog Sound Recordings", AES 20th International Conference, Budapest, 2001.

[3] S. Stotzer, O. Johnsen, F. Bapst, C. Sudan, R. Ingold, "Phonographic Sound Extraction Using Image and Signal Processing", ICASSP, Montréal, 2004.

[4] S. Stotzer, O. Johnsen, F. Bapst, C. Milan, R. Ingold, "Phonographic Sound Extraction using Photography and Signal Processing", Digital Signal Processing Journal, Elsevier, Vol. 17(2), pp. 433-450, 2007.

[5] S. Stotzer, "Phonographic record sound extraction by image processing", PhD Thesis no. 1534, University of Fribourg, 2006.

[6] "Anthology of the Journal of Audio Engineering Society", "Disk Recording" Vol. 1 & 2, 1953-1980.

[7] J. H. Elder, S. W. Zucker, "Local Scale Control for Edge Detection and Blur Estimation", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 699-716, 1998.

[8] S. J. Godsill, P. J. W. Rayner "Digital Audio Restauration", Springer Verlag, 1998.

[9] Phonotechnik, DIN-Taschenbuch 523. Deutsch Institut für Normung, BeuthVerlag, 1991.