SOUND FROM GRAMOPHONE RECORD GROOVE SURFACE ORIENTATION

Baozhong Tian and John L. Barron

Department of Computer Science
University of Western Ontario
London, Ontario, Canada N6A 5B7
{btian,barron}@csd.uwo.ca

ABSTRACT

Preserving historic recording on gramophone records is an important task because the traditional record play back system damages/wears out records eventually. We present an optical flow based method to reproduce the sound signal from gramophone records using 3D robust scene reconstruction of the surface orientation of the walls of the grooves. The imaging setup was modified to overcome a shallow depth of field by using a thin glass plate to obtain additional in-focus parts of the image at a second focal length. The sound signal was recovered from the surface orientation and processed further using the industry standard RIAA filter. The overall algorithm has been tested and found to be working correctly using both undamaged and damaged SP records. The algorithm is a "proof of concept" in that it shows sound can be recovered from time-varying 3D orientation of groove walls.

Index Terms— Gramophone Record Playback via Image Processing, Optical Flow, Surface Reconstruction, Kalman Filtering, Dual Focal Length, RIAA Filtering,

1. INTRODUCTION

There have been a few attempts at developing a non-contact record playback system recently, a detailed literature survey is available [1]. In 2006 we presented an early version of our algorithm and tested it on synthetic data [2]. The simulations showed that the useful frequency response of the system ranged from 200Hz to 20KHz. The Pearson's product correlation coefficient [3] between the sound wave envelops of the original wave and the reproduced wave was about 0.88, which indicated very good reproduction. The groove surfaces of real LP records had to be treated with a thin layer of fine aluminum powder (about 1 μm in diameter), resulting in sufficient texture to facilitate the computation of optical flow.

2. SYSTEM SETUP

The experimental setup shown in Fig. 1 consists of a slowly turning **Turntable**, a **Microscope** that is used to magnify the groove by approximately 400X (40X objective lens, 10X eye piece) to get good images of the groove with its subtle variations and a **Camera**, in our case, a high speed digital camera fitted to the microscope. Images were 24 bit color and were 1280×1024 pixels in size. The records being photographed are 78RPM Standard Play (SP) records. We focus on the 78RPM records since they need less magnification and are easier to illuminate than the 33RPM Long Play (LP) records.

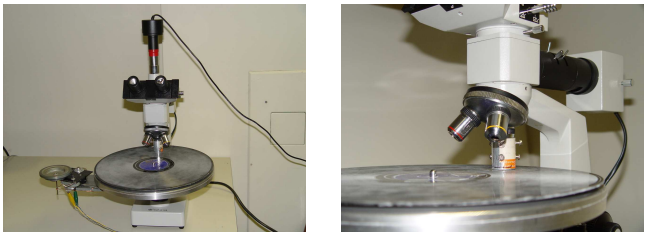


Fig. 1. Setup of turntable, microscope and camera.

3. THE FOCAL LENGTH ISSUE

The recovery of accurate (relative) depth is rather difficult if the target is far away from the camera because the variance in the measurement may exceed the depth difference itself. The average depth of a SP record is about 120 μm. If the average error of depth estimation is $7 - 8\%$, then the average distance between groove and camera should be less or equal to $1.7mm$, which is quite small relative to the focal length of the microscope's objective. In our experimental setup, a 40X objective lens is used to obtain the shortest focal length available. Since a microscope's barrel length is fixed, the microscope is focused by varying the distance of an object from it until the desired image clarity is obtained. The barrel length is the image distance u, which in our case is $u = 160mm$. Because the magnification is 40X, the object distance can be computed as $v = \frac{u}{40} = \frac{160mm}{40} = 4mm$. Then we can compute focal length f using the thin lens equation $f = 1/\left(\frac{1}{u} + \frac{1}{v}\right) = 3.9mm$. This focal length is longer than desired ($1.7mm$), but this is good as our setup allows.

4. THE DUAL-FOCAL MICROSCOPE OBJECTIVE

The main problem with a high power objective lens is the depth of field. For a 40X objective lens, the depth of field is about $1\mu m$ [4]. For a lower magnification factor we obtain a bigger depth of field, but at the same time, the focal length also increases, which makes the difference in depths more difficult to distinguish. For example, a 4X objective lens has a depth of field of about $45\mu m$ [4], which covers about half of the total groove depth. However, the focal length of this 4X objective is about $32mm$. Trying to detect a maximum depth

variation of $120\mu m$ with a focal length of $32mm$ is not feasible. Fig. 2 shows a captured image using the 40X objective.

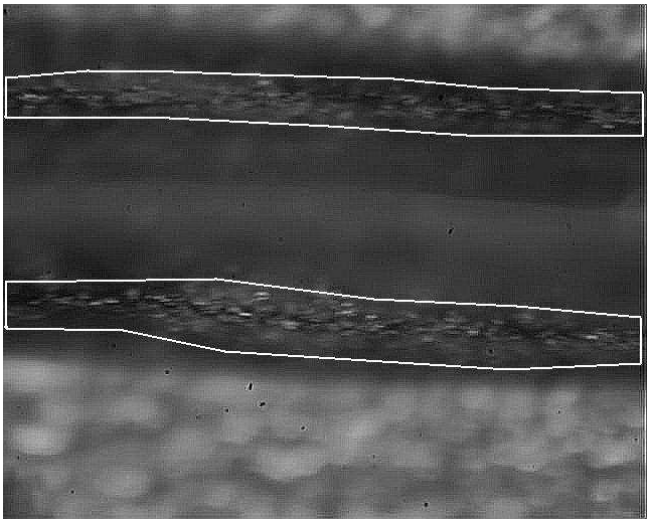


Fig. 2. The groove image captured directly using the 40X objective lens. Notice the shallow depth of field. Only two stripes (shown in highlighted areas) on opposite groove walls of same depth are in focus.

It shows that only two narrow stripes located on the opposite groove walls are in focus.

Since obtaining a larger depth of field, i.e. a clear image all the way from groove top to bottom, is theoretically impossible using our current microscope, bringing more than one depth stripe into focus becomes a more realistic approach. Fig. 3 illustrates the technique of obtaining two focus of

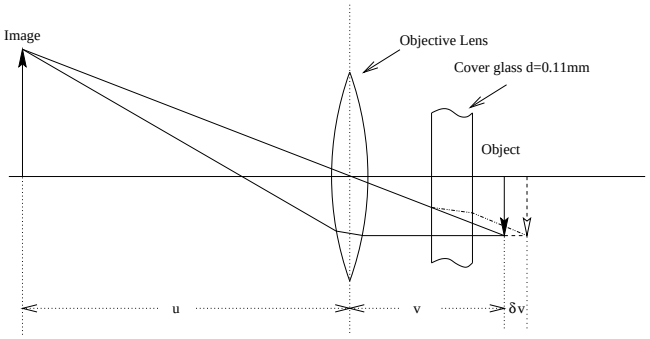


Fig. 3. Illustration of the modified light path for a microscope objective lens, with a piece of thin cover glass inserted between the objective lens and the record surface.

depths using a piece of thin cover glass inserted between the objective lens and the record surface, covering half of the field of view, creating a dual-focusing objective.

When light travels through the piece of cover glass, it is shifted by a certain amount because of the refraction index of glass ($\lambda = 1.515$) is different from that of the air ($\lambda = 1.0$). This shift causes the distance of the in-focused object to increase by a certain amount δv.

Fig. 5a shows a captured image using the modified light path from the same 40X objective lens. Clearly there are three in-focus stripes: the top two are of the same depth, with the groove bottom in the middle, out of focus. The third stripe at the bottom represents a closer stripe on the same groove wall where the second stripe resides. So there are two in-focused regions on the lower groove wall shown in one image. By taking a sequence of such images of the moving groove walls, we can determine their locations by segmenting their optical flow field. After assigning different depth values to these stripe areas, surface orientations are computed based on this information.

5. DETERMINING THE TWO DEPTH LEVELS

Fig. 4 shows the detailed light path, made by enlarging the right half of Fig. 3. From the light incident angle θ and re-

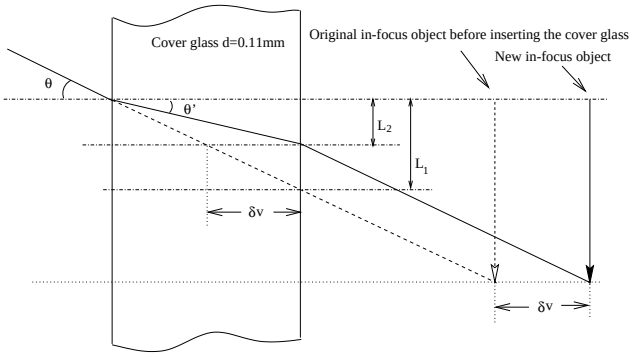


Fig. 4. Illustration of the light path showing the shift of the focus distance of the objective lens when a piece of cover glass is inserted.

fraction angle θ', we have:

$$L_1 = d\,tan(\theta) \text{ and } L_2 = d\,tan(\theta'), \tag{1}$$

where d is the thickness of the cover glass. Then, the shifted focus distance δv can be computed as:

$$\delta v = \left(1 - \frac{1}{\lambda} \cdot \frac{cos(\theta)}{cos(\theta')}\right), \tag{2}$$

where $\lambda = \frac{sin(\theta)}{sin(\theta')} = 1.515$ is the refraction index of the cover glass. The value of $\frac{cos(\theta)}{cos(\theta')}$ is close to 1 when the incident angle is less than $22.5°$. Then Equation (2) becomes:

$$\delta v = d \cdot \left(1 - \frac{1}{\lambda}\right) = 0.037mm, \tag{3}$$

where $d = 0.11mm$. Next, we compute the distances of the bottom two stripes in Fig. 5a using this information. The bottom stripe, since it is not covered by glass. has a distance of $4mm$. The top two are at the same distance of $4.037mm$ using the calculated focus distance shift.

5.1. Optical Flow and In Focus Regions

Optical flow is computed using Horn and Schunck's algorithm [5] modified by adding a groove surface orientation constraint. This constraint regulates the distribution of horizontal optical flow along the vertical direction since it is known

that one of the angles defining wall orientation is $45°$ [1]. Fig. 5b shows the area where optical flow is computed. Optical flow is used to determine the regions which are in focus [out-of-focus regions tended to be heavily blurred with intensity derivatives near zero and poor optical flow], Any location that has an optical flow vector is marked for later segmentation using a Hough transform [6].

5.2. Hough Transform

The computed optical flow for the grooves in Fig. 5a is shown in Fig. 5b and indicates that three stripe regions that are in focus. In order to segment the three regions, we compute a Hough transform using the optical flow position information.

After observing the optical flow positions in many frames, we observed that the stripes are clustered in straight line shapes. At each position that has an optical flow vector, many pairs of ρ and θ values can be calculated according to the normal representation of a line $x\cos\theta + y\sin\theta = \rho$ that passes through this position. At every optical flow position, each value of θ from $-45°$ to $45°$ in steps of $1°$ is used to compute a ρ value and the accumulator array element at $[\theta, \rho]$ is incremented by 1. The Hough transform is shown in Fig. 5c as a $\theta - \rho$ image.

Three global peaks of (θ, ρ) values in the accumulator array are detected, which represent the three lines. Any (θ, ρ) pair that is within a certain range of the three peak (θ, ρ) values is considered to belong to that region. To segment the flow field, we compute (θ, ρ) values for each optical flow position. If it falls in one and only one of the three peak regions, then it belong to that region. If it falls to two or more regions, then this position lies between two regions and usually is caused by noise and is discarded. Only the lower two regions have different depth values. After successful segmentation of the three stripes, the top one is discarded, as shown in Fig. 5d. Also shown in this figure are the centroid lines for the two regions remaining. Void sections are filled with $[x, y]$ computed from peaks (θ, ρ) values using the $\theta - \rho$ equation corresponding to its region. Depth values are then assigned to these two centroid lines to reduce the computational cost when computing surface orientation.

6. COMPUTING SURFACE ORIENTATION

In 2006 we introduced a robust method of estimating the surface orientation given depth values within a small neighbourhood [2]. Now the depth area has been reduced to two virtually parallel lines with two different depth values. Our robust algorithm also works in this case. We use a neighbourhood of 100 pixels about each line for best results. This size choice represents the best compromise between increasing this size (would attenuate high frequencies sounds) or decreasing the size (would enhance noise effects).

7. SOUND FROM SURFACE ORIENTATION

Sound reproduction is as described in [1], except that the sampling rate is different than that used for the synthetic signals. In a synthetic groove, the sampling rate was 220.5kHz. iiIn a real record, the width of view in each captured image of the groove is about $0.5mm$. Dividing by 1280 pixels, we obtain

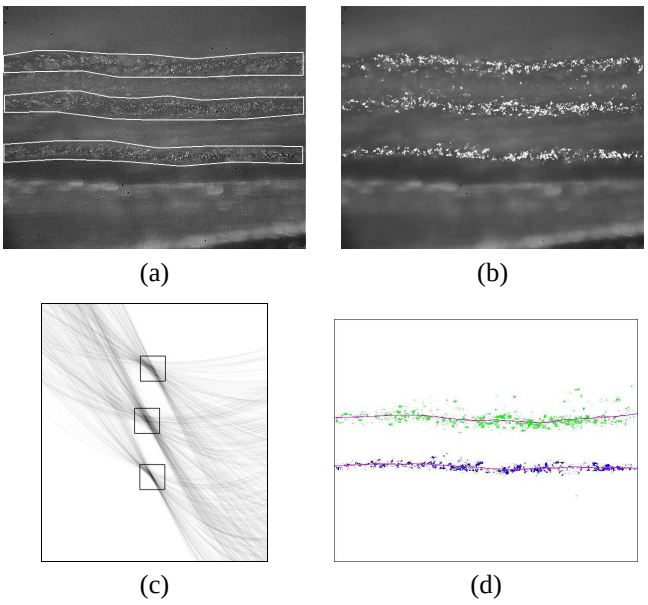


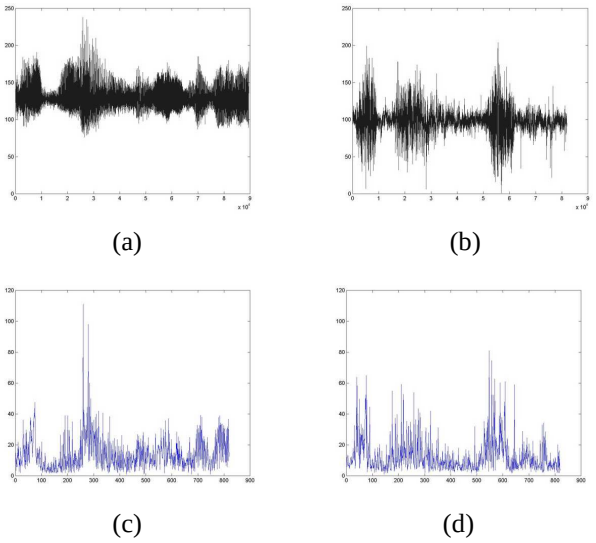
Fig. 5. (a) Groove image captured using 40X objective lens. Notice there are only three stripes (highlighted areas) on opposite groove walls of different depths that are in focus. (b) Locations of computed optical flow (white dots) for this area. (c) Results of Hough transform. The 3 square boxes indicates the 3 peaks that represent the 3 clustered lines. (d) Two segmented regions and their centroid lines.

$\Delta_d = 0.39\mu m$ per pixel. The tangential speed of the groove (at radius=115mm) if played by a 78RPM record player is: $V_t = \omega R = 2\pi 78/60 \cdot 115mm = 939.34mm/s$. The sample frequency at the pixel level is $f_s = \frac{V_t}{\Delta_d} = 2404.71kHz$, which is about 110 times the recording sample frequency of a traditional record player.

There are 6000 groups of such images, with each group containing 12 images for the optical flow computation. It took nearly a week to capture these images and another week to process them (so currently we are nowhere near real-time!!). About 3 seconds of music was reproduced from this effort. Fig. 6 shows the original and reconstructed piece of music from "A Fine Romance" by Kern-Fields. The original music contains both vocal instrumental sounds. The computed Pearson's product-moment correlation coefficient between them is $r = 0.559$. Listening confirms[1] that the vocal sound is quite recognizable in spite of the presence of noise. The music part is lost in the noise, probably because it is already very weak during the part.

Fig. 7a shows the damaged groove from "Give a little, take a little" by Hank Thompson. Nearly half of the groove walls in this image is missing. Fig. 7b shows the positions of the computed optical flow. There are some vertical stripes caused by the damage. The Hough transform segments these lines, ignoring the vertical stripes. Fig. 7c shows the results of Hough transform. Fig. 7d shows the segmented two regions

[1] http://www.csd.uwo.ca/faculty/barron/AUDIOFILES



(a) (b) (c) (d)

Fig. 6. (a) Original sound wave recorded from a 78RPM record player. (b) Recovered sound wave from groups of groove images. (c) Wave peak envelope of the original sound. (d) Wave peak envelope of the recovered sound.

and their centroid lines. The missing parts has been filled in using its the recovered line parameters. Fig. 8 shows the original and reconstructed music. The Pearson's product-moment correlation coefficient between them is $r = 0.564$. Listening confirms that the music sound is quite recognizable in spite of the presence of noise. The popping noise in the original sound has been attenuated in the reconstructed sound.

8. CONCLUSION

We presented an algorithm for reproducing sound from gramophone records using Computer Vision algorithms. A dual-focus setup was used to overcame a shallow depth of field. Future work includes using better optical flow methods, a better imaging setup, a better depth/orientation reconstruction, and a parallel (SIMD) implementation to make the algorithm real-time.

9. REFERENCES

[1] Baozhong Tian, *Reproduction of Sound Signal from Gramophone Records using 3D Scene Reconstruction*, Ph.D. thesis, University of Western Ontario, 2006.

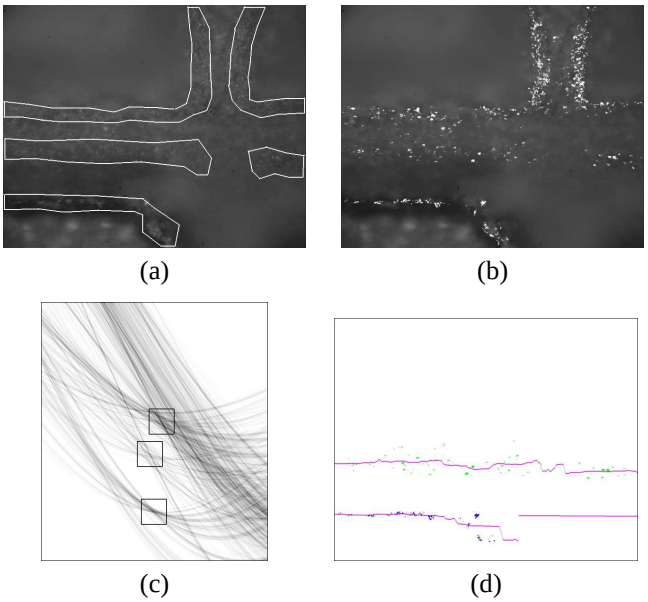
[2] B. Tian and J. L. Barron, "Reproduction of sound signal from gramophone records using 3d scene reconstruction," in *Irish Machine Vision and Image Processing Conference*, Dublin, Ireland, August 30th - September 1st 2006, pp. 84–91.

[3] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, 2 edition, 1992.

[4] Nikon, "Nikon microscopy tutorial," Internet reference: www.microscopyu.com, 2006.

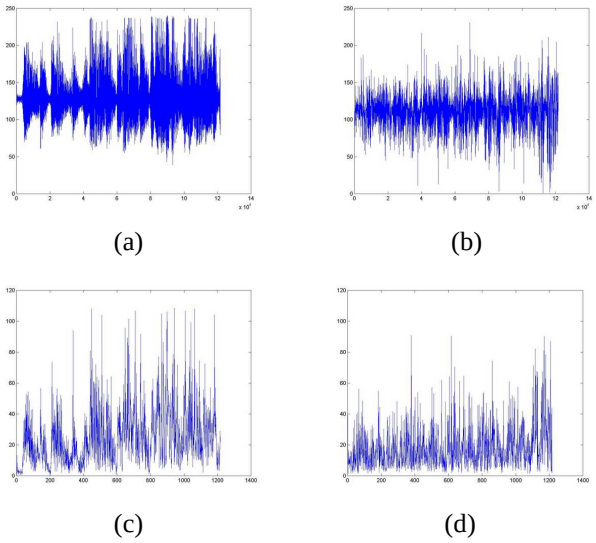
[5] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–204, 1981.

[6] Richard E. Woods Rafeal C. Gonzalez, *Digital Image Processing*, Addison-Wesley Publishing Company, 1992.



(a) (b) (c) (d)

Fig. 7. (a) An image showing a damaged groove. (b) Positions of the computed optical flow. (c) Results of Hough transform. The 3 square boxes indicate the 3 clustered line regions. (d) Two segmented regions and their centroid lines. The missing parts has been filled in using the computed line parameters.



(a) (b) (c) (d)

Fig. 8. (a) Original sound wave recorded from a 78RPM record player with audible pops. (b) Recovered sound wave from groups of groove images. (c) Wave peak envelope of the original sound. (d) Wave peak envelope of the recovered sound.