Reproduction of Analog Record Sound Using Digital Image Captured by a Flatbed Scanner
- Comparison of Sound Groove Edge-Extraction Filter for Stereo Records -[1]

SHIKAMA, Shinsuke[2]　Department of Electrical and Electronic Engineering, Setsunan University
MIZUNO, Hiroaki　Graduate School of Science and Engineering, Setsunan University

Abstract

We have studied on a non-contact reproduction method of sound signal from phonograph records based on digital image processing. First, we examined whether a groove geometry of stereo-record could be digitized based on a resolution of a flatbed scanner which was commercially available. Next, we investigated three filtering methods to extract the groove edges. As a result, we found that Laplacian of Gaussian (LoG) and Difference of Gaussians (DoG) filter showed relatively superior edge extraction characteristics. From the filtered image, we digitized sound signal. The reproduced sound contained strong high-frequency noise superimposed on the original sinusoidal wave. Among the three filters, the DoG filter showed most preferable time waveform and the LoG filter showed minimum harmonics distortion components. From these results, we have concluded that the extraction accuracy of the groove edges should be improved further in order to reduce the disturbance of the filtered waveform.

Keywords : analog record, flatbed scanner, digital image processing, edge extraction, sound signal

1. Introduction

Currently, production volume of the analog record becomes the minimal number and becomes valuable. Since sound information is directly minced on the record disk, it may become impossible to reproduce the original sound by the wound or crack which stuck to a sound track by a certain cause. Furthermore, it is a big problem that a sound track is gradually worn by repeating normal reproduction and the deterioration of the record disk progresses even if we use a superior record player.

To solve such a problem, studies aiming at non-contact reproduction of phonograph records have been reported from multiple institutions. Among those, the laser turntable devised by Stoddard et al. of Stanford University is a practical example[(1), (2)]. This is intended to realize a non-contact playback by detecting the

[1] 【原稿受付】2016 年 7 月 28 日、【掲載決定】2016 年 10 月 6 日
[2] 【主著者連絡先】 鹿間　信介　　摂南大学、准教授　　e-mail: shikama@ele.setsunan.ac.jp
〒572-8508　大阪府寝屋川市池田中町 17-8、摂南大学理工学部 電気電子工学科

sound information from the reflected light obtained by irradiating the sound track using laser beam. Products using this principle is manufactured and sold by ELP Corporation[3]. But they are expensive products, it can not be said to be common.

Aside from this, it was reported that Uozumi of Hokkai Gakuen University used a dedicated optics composed of optical microscope and CCD camera, acquired and processed the sound track image of the monophonic Short-Play (SP) records, and succeeded to reproduce the sound signal[4]. Johnsen et al. of University of Fribourg reported a playback method of monophonic 78rpm SP record by a hybrid system using an analog film and a linear CCD sensor[5]. They took an analog picture of a disk in order to preserve the sound information. The film was digitized using a special rotating scanner equipped with the linear CCD sensor mounted on microscope optics. The sound was then extracted from the digital image by measuring the radial displacement of the groove. Their system takes many steps to reproduce the sound information.

In this study, we captured the sound track image of the stereo record using a flatbed scanner for PCs, and examined how to reproduce record sound using digital image processing. The flow of the whole processing is shown in Fig. 1. Comparatively early researches using image scanner are disclosed on the web by Springer[6], and Olsson et al[7]. Each of these aims at reproducing the sound of the monophonic SP records. Since the principle of the former is indefinite, the difference from our research cannot be shown. The latter is extracting monaural sound by performing data processing for detecting displacement of the sound track center on the disc plane. In this study, we aimed to reproduce audio information by applying appropriate edge-extraction filter for stereo records in the pre-processing stage in Fig. 1. Because most of analog records are black, captured digital image contrast is low. So it was compared and examined by the point of view of noise reduction of filtered image and good sound reproduction capability about three kinds of edge-extraction filter.

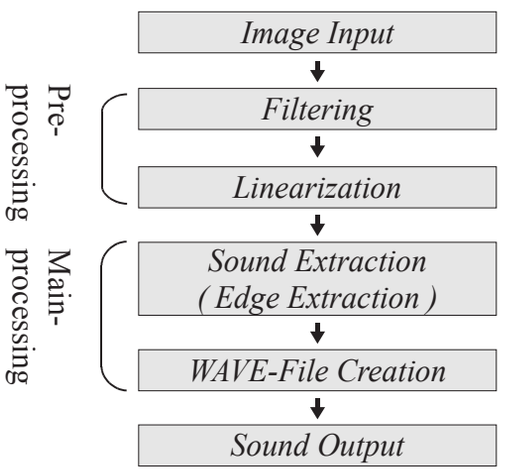


Fig. 1 Processing overview.

2. Image Capturing

2-1 Standards of records [8],[9]

Currently, stereo sound information is engraved on the record disk in the standard method of 45-45 stereo system. According to this system, the independency of the sound information on a right-and-left channel is secured by the information on the left channel being engraved on the wall of the direction of record inner circumference, and the information on the right channel being engraved on

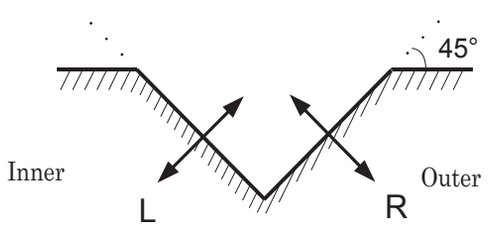


Fig. 2 Sound groove (45-45 stereo system).

the wall of the direction of the record perimeter (see Fig. 2).

The sound track of a monophonic record has constant groove width, and it is determined as 50μm or more by the Japanese Industrial Standards (JIS). On the other hand, in a stereo record, groove width is not fixed. When it is 30cm of 33 1/3rpm disk, the groove width changes between 40 and 90μm, and when it is 17cm of 45rpm disk, the groove width changes between 60 and 120μm. Below, the amplitude limit of sound groove is explained.

It is the conditions of the amplitude limit of a sound groove that the adjoining sound grooves do not have an intersection (overload). If the average pitch of a sound track is set to P as shown in Fig. 3, and the permitted minimum width of a sound groove is set to W, the amount "a" of the maximum deviation of the sound groove to a slope normal direction is given by:

$$a = \frac{1}{4\sqrt{2}}(P - W). \tag{1}$$

Moreover, even if the frequency of the audio signal is the same, in the perimeter and inner circumference, the record wavelengths of a sound track differ: a wavelength becomes shortest at the most inner circumference. Let L be the track running distance per one second, f be the frequency of the audio signal, r be the radius of a sound track, N be the number of revolutions of the record per minute, and then record wavelength λ of the sound track is given by Eq. (2).

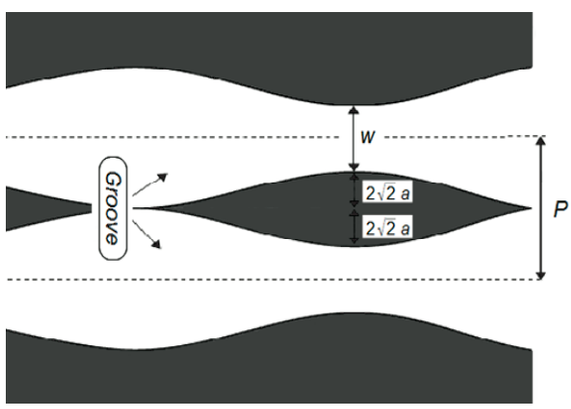


Fig. 3 Amplitude shift limit of stereo records.

$$\lambda = \frac{L}{f} = \frac{2\pi rN}{60f}. \tag{2}$$

For reference, various dimensions specifications of records are shown in Table 1.

Table 1 Dimensions specifications of records (mm).

Size	**17 cm**		**30 cm**
Rotational Speed	**45 rpm**	**33 1/3 rpm**	**33 1/3 rpm**
Outside Diameter	175±1	175±1	301±2
Diameter of Outermost Groove	168 max	168 max	293 max
Minimum Inside Diameter of Recording	106 min	106 min	115.2 min
Stopping Groove Diameter	97±1	97±1	106.4±1
Eccentricity Allowance (hole)	−	−	0.2 max

2-2 Resolution of scanner

In this research, a flatbed scanner for PC (GT-X820 made by EPSON) was used as an image input device. Since the maximum resolution of the main scanning direction of this scanner is 6400dpi, a sound

track image is sampled at intervals of 3.97μm. If the minimum width of the sound track of a record to be W=40μm and the average value of the sound track pitch by real measurement to be P=128μm, the peak magnitude of a=16μm is calculated by using Eq. (1). Moreover, by the plane image of the sound groove edge locus acquired with the scanner, the groove amplitude to the normal direction of the slope which is inclined 45 degrees is enlarged to $\sqrt{2}$ times. From the above consideration, acquisition of a sound groove image can be performed satisfactorily. In addition, although there are restrictions of a dynamic range, it is expected that sound information being reproduced using a scanner acquired images.

Next, let us think on the record wavelength of a sound groove. It will become shortest when we record the sound of the highest frequency on the most inner part of the disk rotating at 33 1/3rpm. If maximum audible frequency of human to be f=20kHz and the track radius of the most inner part to be r=53mm from Table 1, the minimum wavelength of λ_{min} recorded on the disk is estimated to be 9.25μm from Eq. (2). Therefore, the sampling theorem is satisfied also in a direction along the track and the sound signal recorded on the disk is theoretically reproducible.

2-3 Image input

Stereo record of 17cm 33 1/3rpm[10] and "Frequency Record" of 30cm 33 1/3rpm[11] which was supplied for evaluating audio equipments were used in this study. About the former, the scanner performed sound track picture acquisition of the whole disk. A predetermined portion image of the latter was used as a single frequency signal source for sound reproducibility experiments. First, the image of 17cm disk was acquired with the scanner controlled by a scanner driver. In order to obtain the image of the whole disk in the resolution of 6400dpi, we divided the image into eight sub-areas and downloaded each image-data into the personal computer in the form of 8 bit gray scale picture (299 MB per division) of BMP format (Fig. 4). We could carry out picture acquisition of the 17cm record only by 8 division method. The reason for this was by the specifications of the data capacity of the scanner driver used in the highest resolution mode. This method acquired the image without moving the record itself. Therefore, no problem occurred on the connection of the groove images on each sub-area.

Fig. 4 Image capture of a record divided into eight sub-areas.

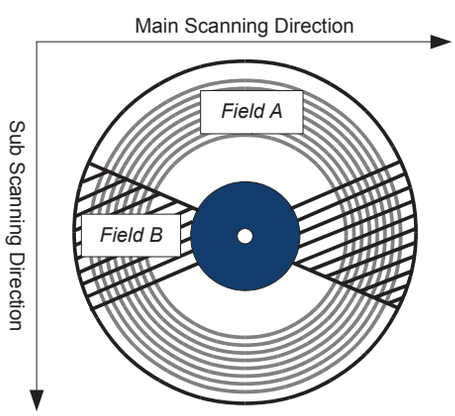


Fig. 5 Image of a scanned record.
A: Sound groove field in good condition.
B: Sound groove field in the form of shadow.

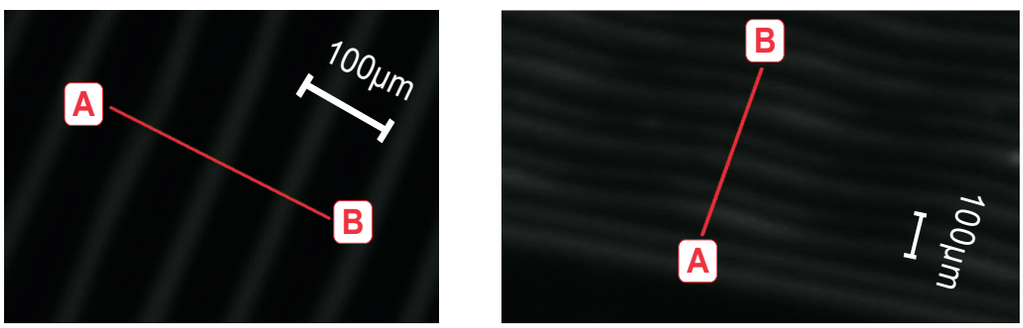
3. Pre-processing

Before extracting sound information by the

following main-processing, the groove image was processed into the optimal state by the image processing described in this chapter. Especially we experimentally compared various image filtering methods in order to extract the clear edge locus of the sound groove including sound information.

3-1 Image filtering

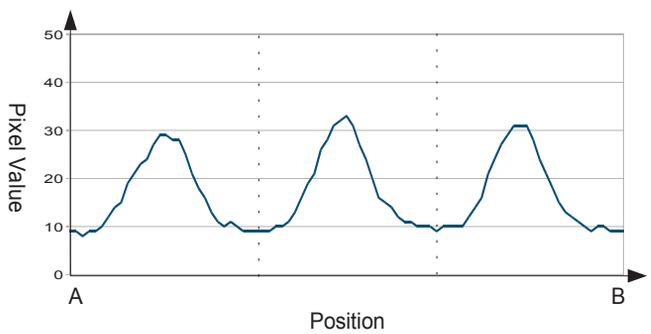
When acquiring an image of an entire disk by the scanner, the experiment showed that a gloomy shadow area appeared in the specific angular zone of the whole image. This occurred with the difference in the



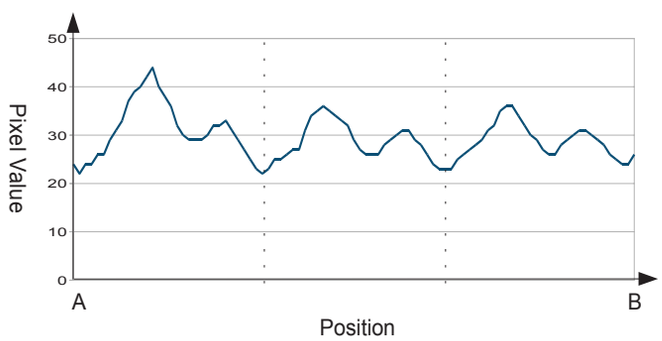
(a) Image 1 (b) Image 2

Fig. 6 Grayscale image in arbitrary position of Field-A: Red line shows the extent between A and B. "A" is on the inner side of the disk, "B" is on the outer side of the disk.

surface reflection characteristic between each record. Also the illumination characteristics of the sound track using a linear illumination lamp provided on the scanner-head was thought to be the cause of the phenomena. To be more specific, the Field-B of the sound track shown with hatching in Fig. 5 was the shadow-like area. And it became difficult to extract the clear groove edges within the area. In the Field-B, tracks make almost right-angle to the linear light source which was along the main scanning direction. In this paper, we applied filter processing to the sound groove images in the Field-A, and tried to get a better result. Light was uniformly illuminated on the sound groove in the Image-1 of Fig. 6(a) which was acquired at an arbitrary portion in the Field-A (disk: 17cm 33 1/3rpm). Therefore, the change of



(a) Image 1



(b) Image 2

Fig. 7 Pixel value waveform between A and B in Fig. 6.

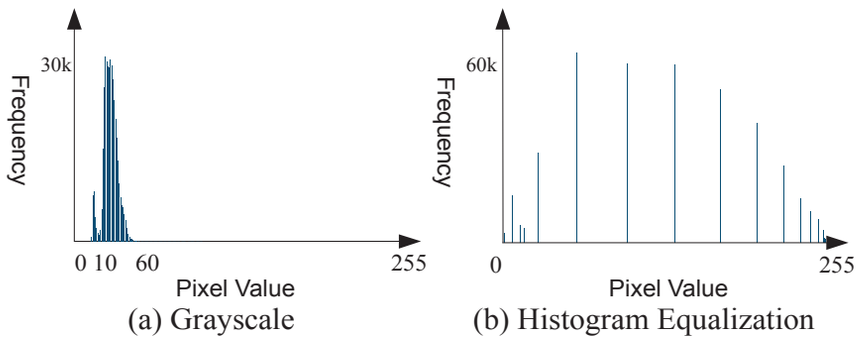
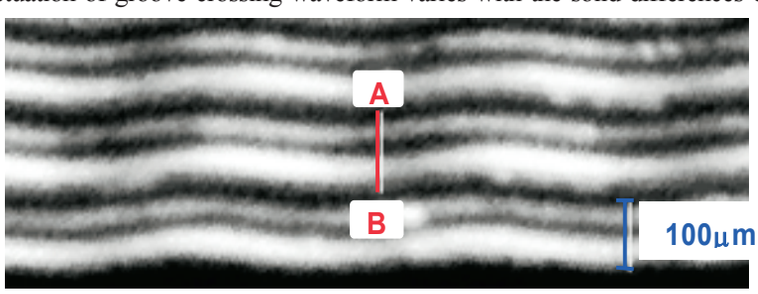
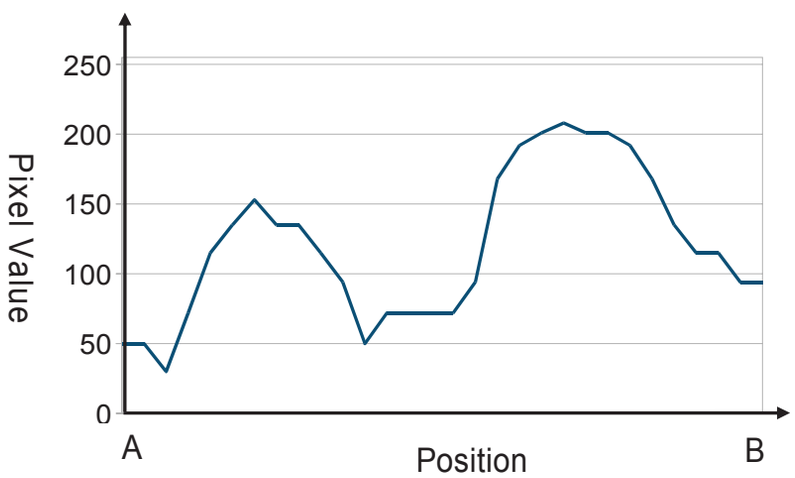


Fig. 8 Histogram of "Image 2" in Fig. 6(b).: Pixel value 0 corresponds to black, 255 corresponds to white.

a pixel value waveform along the line A-B shown in Fig. 6(a) was periodic, and fluctuation of amplitude and a local average level was small (see Fig. 7(a)). Here, 'A' was positioned on the inner side of the disk, and 'B' was on the outer side of the disk. In such an image, it was considered possible to detect excellent edges using binarization by setting a threshold value of 20 or so. However, as a sound groove image like the Image-2 of Fig. 6(b), which was acquired at an arbitrary portion in the Field-A (disk: 30cm 33 1/3rpm), fluctuation of the peaks and valleys of a pixel value waveform along A-B and that of local average level was comparatively large, and a threshold cannot be defined appropriately (see Fig. 7(b)). In addition, it turned out that degree of fluctuation of groove-crossing waveform varies with the solid differences of records, and it



(a) Gray scale image



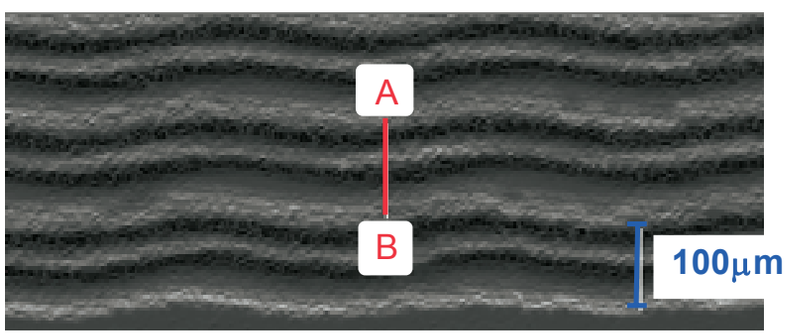
(b) Pixel value waveform between A and B

Fig. 9 Histogram equalization processing result on Fig. 6(b). (a) Gray scale image; Red line shows the extent between A and B. "A" is on the inner side of the disk, "B" is on the outer side of the disk. (b) Pixel value waveform between A and B (i.e. one cycle of the sound groove).

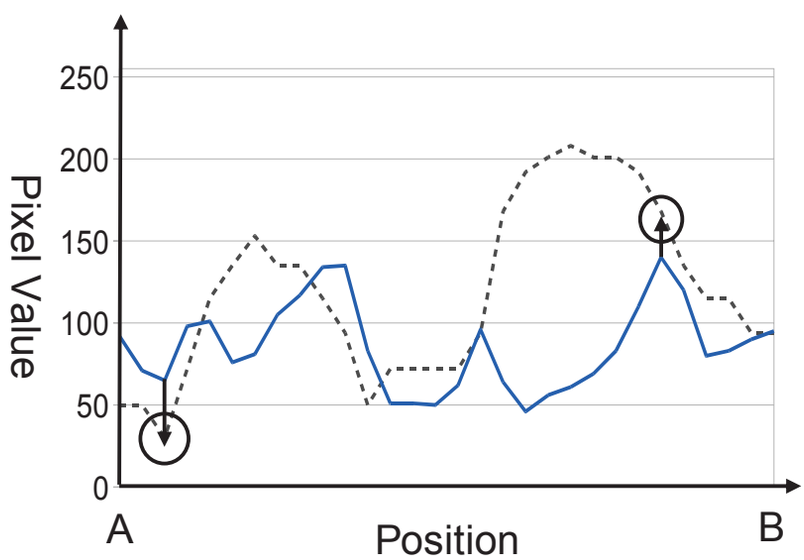
also varied greatly with the differences of the angular position even within the same record. Pixel values of Fig. 6(b) were concentrated in the region of 10-60, as shown in the histogram (Fig. 8(a)). This made the image shown in Fig. 6(b) low contrast and inferior in visibility. Then, we equalized the histogram[12] making original pixel value distribution expand to the range of 0-255 (Fig. 8(b)). As a result, the contrast of the sound-groove crossing waveform was improved, as shown in Fig. 9(b). The improvement of the visibility of the sound groove image was also significant (Fig. 9(a)). Sound information was recorded as the displacement of the sound groove edges (i.e. undulation of the grooves). Therefore it is important to accurately detect the sound groove edges. In the following sections, we examine three edge-extraction methods to the image which contrast is improved by the histogram equalization. Incidentally, in subsequent image processing, compensations are made for preventing information loss caused by minus components of pixel values. In other words, by adding a constant offset to the entire image, the pixel values of the edge extracted image are set to be positive.

(1) Sobel filter

First, we tried the Sobel filter processing[13]. It performed a smoothing process in the horizontal 3 pixels



(a) Gray scale image.



(b) Pixel value waveform between A and B.

Fig. 10 Sobel filter processing result on Fig. 9(a). (a) Gray scale image.; "A" is on the inner side of the disk, "B" is on the outer side of the disk. (b) Pixel value waveform between A and B (i.e. one cycle of the sound groove).

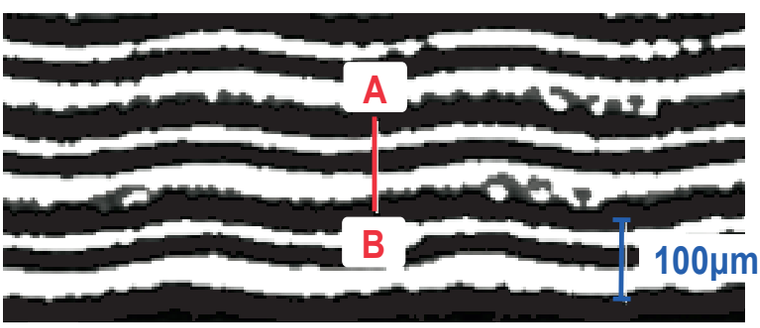
(weighted in the center pixel) and differentiation in the vertical direction by a 3×3 kernel. Processing with a kernel that was rotated 90 degrees was combined at the same time. The filtered image and waveform are shown in Fig. 10. In this case, edges can be read at the overlapped points of peak or valley of the Sobel-filterd waveform and the histogram-equalized waveform in Fig. 10(b). It was considered that the noise components superimposed on the Sobel-filtered waveform were reduced by the effect of smoothing.

(2) Laplacian of Gaussian (LoG) filter

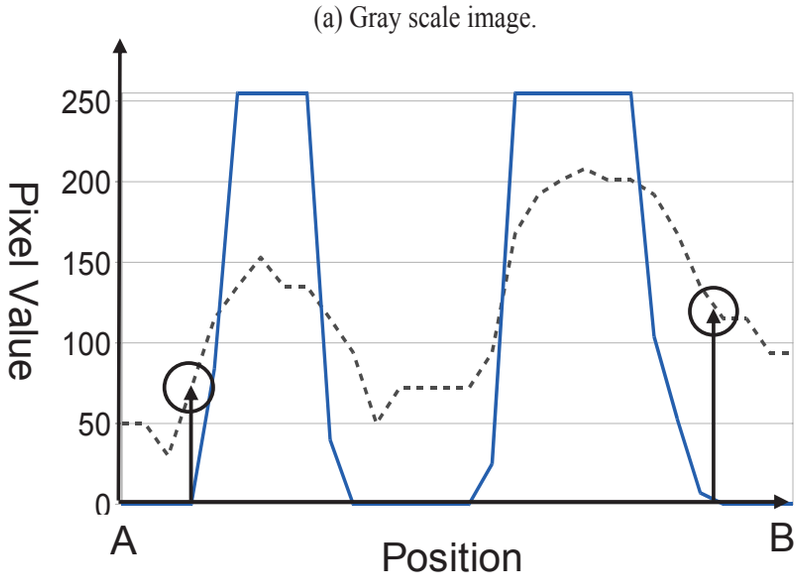
Then we tried the LoG filter [14]. Laplacian was applied to an image that has been smoothed with a Gaussian filter in order to reduce the image noise that occurred in the second derivative of the Laplacian. Filter coefficients were calculated by the following equations:

$$G_\sigma(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{3}$$

$$LoG\equiv\frac{\partial^2}{\partial x^2}G_\sigma(x,y)+\frac{\partial^2}{\partial y^2}G_\sigma(x,y)=\frac{x^2+y^2-2\sigma^2}{2\pi\sigma^6}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{4}$$



(a) Gray scale image.



(b) Pixel value waveform between A and B.

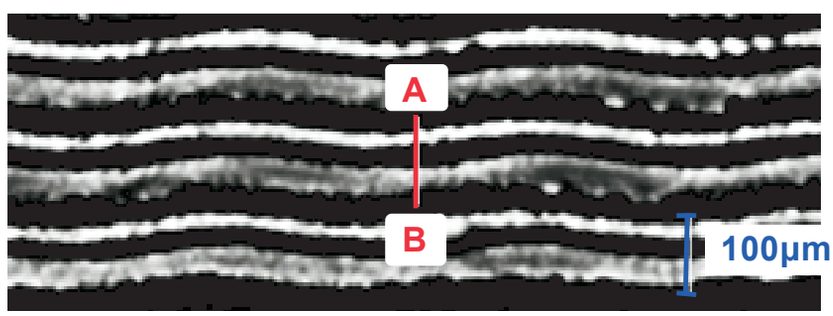
Fig. 11 LoG filter processing result on Fig. 9(a). In Eq. (4), we set σ=2, and applied kernel size of 7×7. (a) Gray scale image.; "A" is on the inner side of the disk, "B" is on the outer side of the disk. (b) Pixel value waveform between A and B (i.e. one cycle of the sound groove).

Where, (x, y) are the image coordinates and σ is a standard deviation of the Gaussian function $G_{\sigma}(x, y)$.

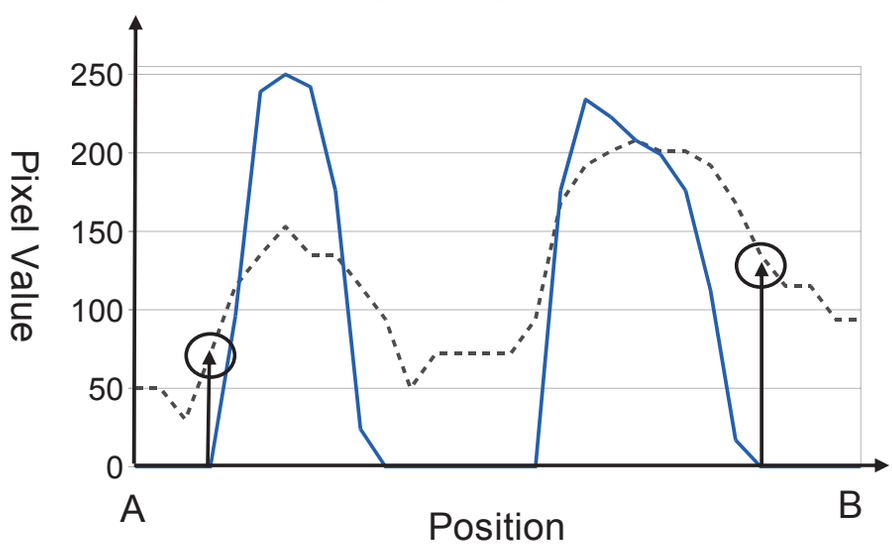
The filtered image and waveform are shown in Fig. 11. As can be seen in Fig. 11(a), the groove image was changed to be mostly binarized, and fine noise components seen in the Sobel filtered image (Fig. 10(a)) were removed. Since Laplacian was secondary differentiation, the points of the filtered waveform crossing zero were considered to be the groove edges. Therefore, it was thought that the points shown by the circles in Fig. 11(b) corresponded to the groove edges of an inner and outer circumference respectively.

(3) Difference of Gaussians (DoG) filter

The DoG filter[14] removes high frequency detail that often includes random noise, and it is used as one of the most suitable filter for processing images with a high degree of noise. Filter coefficients were calculated by the following equation: the difference of Gaussian functions which had the two standard deviations of σ_1 and σ_2.



(a) Gray scale image.



(b) Pixel value waveform between A and B.

Fig. 12 DoG filter processing result on Fig. 9(a). In Eq. (5), we set σ_1=3, σ_2=1, and applied kernel size of 11×11. (a) Gray scale image.; "A" is on the inner side of the disk, "B" is on the outer side of the disk. (b) Pixel value waveform between A and B (i.e. one cycle of the sound groove).

$$DoG \equiv G_{\sigma_1}(x,y) - G_{\sigma_2}(x,y) = \frac{1}{2\pi\sigma_1^2}\exp\left(-\frac{x^2+y^2}{2\sigma_1{}^2}\right) - \frac{1}{2\pi\sigma_2^2}\exp\left(-\frac{x^2+y^2}{2\sigma_2{}^2}\right). \quad (5)$$

If the parameters σ_1 and σ_2 are close to the limit, the DoG is known to become the same as the LoG except for the proportionality coefficient[14]. Therefore, by setting two parameters (σ_1 and σ_2) of the DoG function, we could increase the flexibility of characteristic control of edge extraction, while performing noise suppression of the groove image using Gaussian functions. The filtered image and waveform are shown in Fig. 12. The Fig. 12(a) shows that the DoG filter removed the fine noise components seen in the case of the Sobel filter (Fig. 10(a)). This was similar to the processing result using the LoG filter. Because DoG function was approximated by secondary differentiation of Gaussian, the points of the filtered waveform crossing zero were also considered to be the groove edges: The points shown by the circles in Fig. 12(b) correspond to the groove edges of an inner and outer circumference respectively.

Comparisons of the above three kinds of filtering processing results are summarized in Table 2.

Table 2 Comparison of filtering methods.

Filter		Sobel	LoG	Dog
Result	Noise	Less	Minimum	Minimum
	Edge Detection	Possible	Possible	Possible
Evaluation		Good	Very Good	Very Good

3-2 Linearization

By linearizing a pseudo-circular sound track, subsequent signal processing to extract the waveform of the sound becomes easy. In this section, we describe the concept of the algorithm that we adopted with respect to linearization.

(1) Calculation of central point coordinates and a radius

The central point coordinates and radius of a sound track locus served as basis of linearization. Therefore, we created a program which calculates these parameters along the following ideas.

(a) A temporary center point was decided by eye measurement.

(b) We run a large number of search line toward the temporary center point from outside of the record disk.

(c) We presumed a disk contour from the edge points $(x_a, y_a), (x_b, y_b)$, which were found on the search line started from outside of a record (Fig. 13).

(d) Center coordinates (x_c, y_c) and radius r were calculated using the points on the presumed disk contour. The algorithm was as

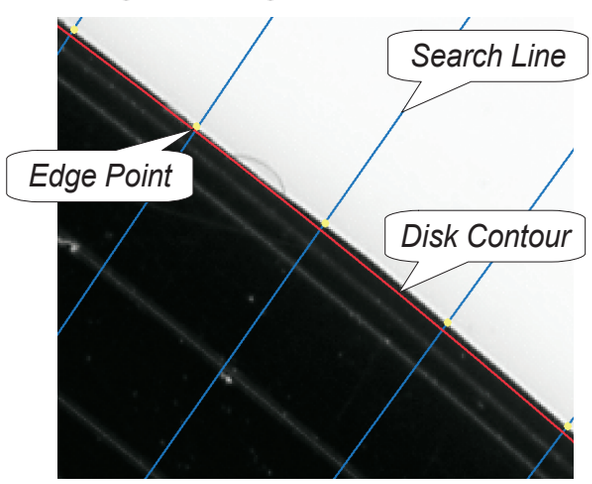


Fig. 13 Detection of the edge points on disk contour.

follows.

First, the theorem "ABC will become a right-triangle if the diameter of a circle is set to AB and arbitrary points C other than A and B are taken on this circumference" is known well (Thales' Semicircle Inscribed Angle Theorem[15]). Using this theorem, the plural right-triangles (*K* pieces) were drawn on the perimeter circle of a record disk (Fig. 14). And by averaging the coordinate's values of hypotenuse end points, the central point coordinates (x_c, y_c) are calculated as follows:

$$\begin{cases} x_c = \dfrac{1}{K}\sum_{i=1}^{K}\dfrac{x_{ai}+x_{bi}}{2} \\ y_c = \dfrac{1}{K}\sum_{i=1}^{K}\dfrac{y_{ai}+y_{bi}}{2} \end{cases} \tag{6}$$

Here, (x_{ai}, y_{ai}), (x_{bi}, y_{bi}) are the coordinates of the i-th diameter's end points. By using (x_c, y_c) obtained in Eq. (6), radius *r* of the disk is given by the average operation:

$$r = \frac{1}{K}\sum_{i}^{K}\sqrt{(x_{ai}-x_c)^2+(y_{ai}-y_c)^2}\ . \tag{7}$$

When five disk images were actually acquired using the same 17cm record and the radius was presumed by the above algorithm, it turned out that measured values are distributed over the range of ±21μm from average value. Compared to eccentricity allowance of the center hole 0.2mm (standard value in a 30cm record, no standard value of a 17cm record specified) as shown in Table 1, which was a value smaller by one digit. Therefore, it was considered that the central point coordinates (x_c, y_c) and the radius *r*, which were calculated by using the above-mentioned algorithm, could be satisfactorily used for the following coordinate transformation processing.

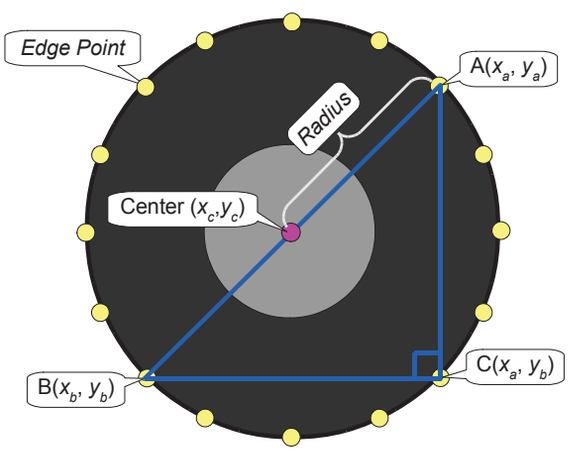


Fig. 14 Calculation of central coordinates and radius.

(2) Coordinate transformation

In order to linearize a circular sound track locus, coordinate transformation of the image was carried out. Here, the sound track trajectory was a spiral originally. It was assumed that could be regarded as concentric because spacing of the grooves was very narrow. As shown in Fig. 15, let (x_s, y_s) be the coordinates before transformation, and (x_c, y_c) be the center point coordinates, *r* be the radius, then the following relationship holds:

$$(x_s, y_s) = (r_j \cos\theta_i - x_c,\ r_j \sin\theta_i - y_c)\ . \tag{8}$$

Where,

$$r_j = r - j\times\Delta r \quad (j = 0,1,\cdots,m), \tag{9}$$

$$\theta_i = i\times\Delta\theta \quad (i = 0,1,\cdots,n), \tag{10}$$

$$\Delta r = r/m, \tag{11}$$

$$\Delta\theta = \pi/n. \tag{12}$$

Then, the following equation gives the linearized coordinates (x_o, y_o) :

$$(x_o, y_o) = \left(\frac{\pi(n-2i)}{2n} r_j - x_c,\ r_j - y_c \right). \tag{13}$$

And after performing interpolation processing using the pixel value near the coordinate (x_s, y_s), linearization was completed by carrying out resampling as pixel data of grid point indicated by Eq. (13).

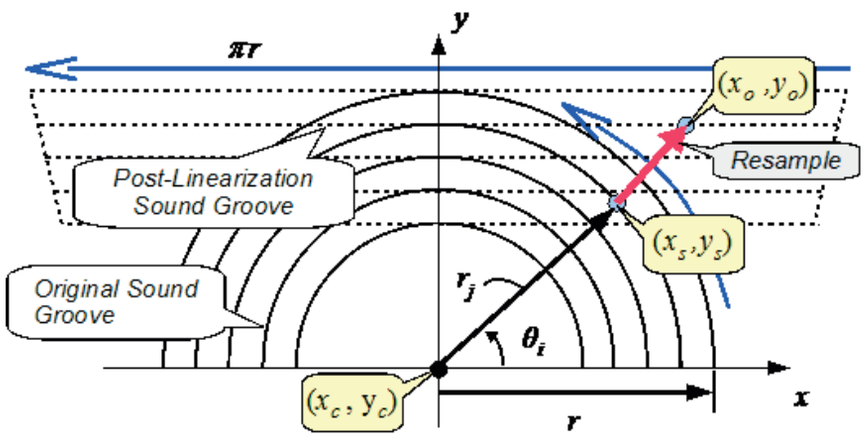


Fig. 15 Linearization using coordinate transformation.

4. Sound Extraction

In order to restore the signal waveform stored on a record disk, it was necessary to recognize the placement of each sound groove trajectory that had been linearized from the original image. Prior to this, we had to define the domain (ROI: Region of Interest) of a signal to extract. Here, we used the images which performed each filter processing of Sobel, DoG, and LoG, from the result of the previous chapter.

4-1 Detection of a sound track edge position

In order to decide ROI, the position of each sound track locus was detected. First, many search lines were run to y-direction within the limits of the track image. Both the edges of a sound track were detected by the pixel value change using a certain threshold on the search line. Moreover, the sound track range on a

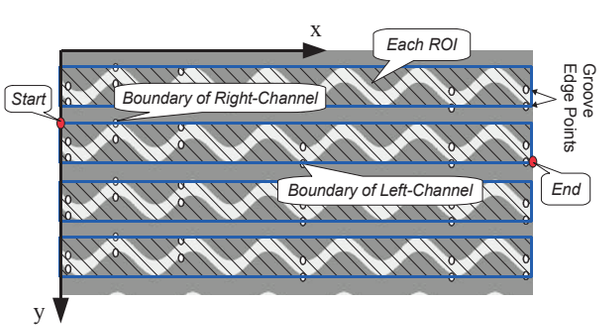


Fig. 16 Setting of ROI.

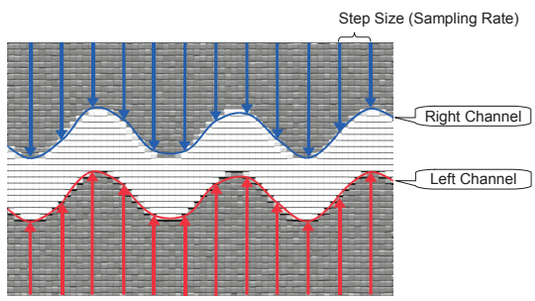


Fig. 17 Sampling of sound waveform.

record was decided by the standard (JIS S 8502[8]: now obsolete). It was determined according to this standard whether the detected edge points were within the range of a valid sound track.

4-2 Determination of ROI

From the edge points detected by the method described in the previous section, we determined the ROI rectangular (Fig. 16). The starting point of the ROI had a minimum y-coordinate value of the right channel, and its x-coordinate was 0. In addition, the end point of the ROI had a maximum y-coordinate value of the left channel, and a maximum x-coordinate value.

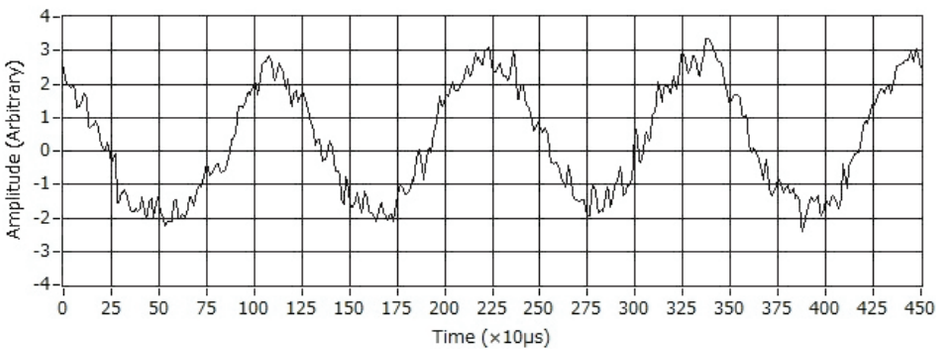
4-3 Sampling and quantization

In order to extract a stereo sound signal waveform from the sound grooves in each ROI, the left-channel was sampled by scanning at arbitrary intervals (step size) upwards from the bottom. And the right-channel was scanned and sampled downwards from the top within the ROI (Fig. 17). More specifically, left-channel waveform was detected by the coordinates of bottom value which appeared first on the scanning lines in the Sobel filtered image, and right-channel waveform was detected by the coordinates of peak value which appeared first on the scanning lines (Fig. 10(b)). And we made those y-coordinates the amplitude values of sound signal waveforms. Moreover, in the pixel value waveform which performed LoG/DoG filter processing, the zero crossing points which appeared first on the scanning lines were detected, and we made y-coordinate values of these points the amplitude values of the sound waveforms of the left-channel and the right-channel (Figs. 11(b) and 12(b)).

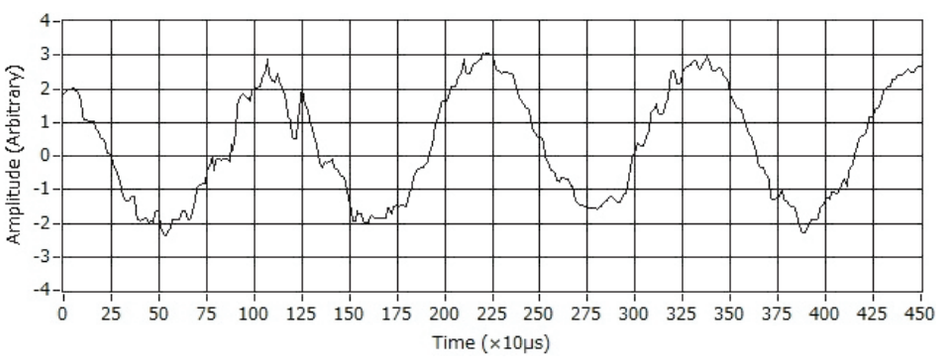
5. WAVE-file Creation and Evaluation

Sound form (i.e. sampling rate, number of bits) was decided based on the waveform data extracted in Chapter 4, and audio file was created as a WAVE-file. The waveforms of 1 kHz sound of the left-channel recorded on $r \approx 75$mm on the "Frequency Record[11]" extracted by our methods are shown in Fig. 18. We digitized the sound waveform with the smallest pixel pitch of 3.97μm both in the x and y directions. This means that the sound groove having wavelength of 262μm was sampled by 66 points per wavelength along the linearized groove. At the present stage, about 11 sound grooves of 1 kHz waveform located in Field-A (Fig. 5) are reproducible. However, it has yet to play a continuous sound from the grooves in the regions of A and B. High-frequency noise components were superimposed on each sine waves of 1 kHz as shown in Fig. 18. The noise components were less in the LoG and DoG filleted signal than the noise in the Sobel filtered signal. In particular, the sound waveform using the DoG filter was thought to be the best because it showed smooth and less high frequency components. The main cause of the noise components were considered as follows: As a result of generating sound waveforms using the edge points presumed by the filtered image (Sobel, LoG, DoG), the detected edge positions were disturbed. The sound of 1 kHz without a feeling of noise was obtained in reproduction by a conventional record player (DP-200USB made by DENON) using a needle type cartridge. The player's sound was encoded into an MP3 format having bit-rate of 192kbps, which was thought be sufficient for reproducing the audio signal up to 20kHz. The spectrum data of reproduced sound are shown in Fig. 19: Figure 19(a) is the sound spectrum reproduced by the conventional record player, and Figs. 19(b), 19(c), and 19(d) are the spectrum of the reproduced sound using

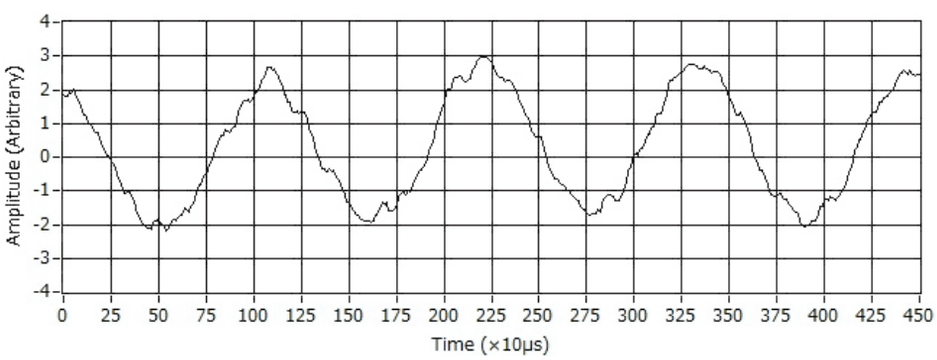
the image processing of this work (Sobel, LoG, DoG). Using the conventional record player, harmonic components to the fundamental at 2 kHz and 3 kHz were -40 dB and -50 dB respectively. According to the result used the Sobel filter, each harmonic component (2 kHz and 3 kHz) was inferior with -30dB and -30dB. In addition, the harmonic components were observed to be -50dB and -40dB when the LoG filter was applied. And they were -40dB and -40dB when the DoG filter was used. The harmonic components using LoG and DoG filter were better than the result using the Sobel filter. In particular, the LoG filtered result was the best among the three methods. We show the summary of measured harmonic components in Table 3. Playback by the WAVE-files (114k samples/sec, 16 bits/sample) transformed from the extracted waveforms (Fig. 18) were perceived noisy and they were auditory different from monotonous and clear sinusoidal sound. We found that improvement of edge detection accuracy was an important issue from these results.



(a) Sobel filter.

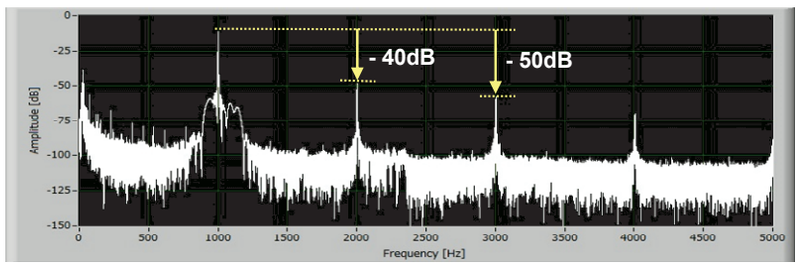


(b) LoG filter.

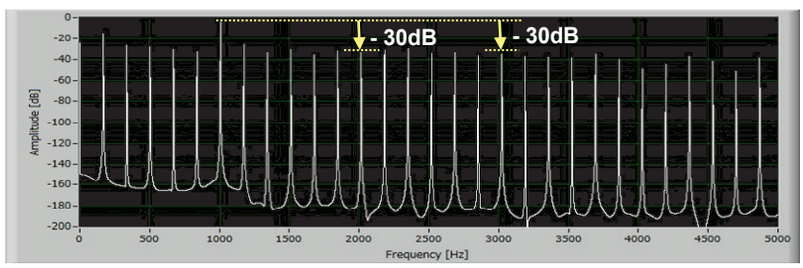


(c) DoG filter.

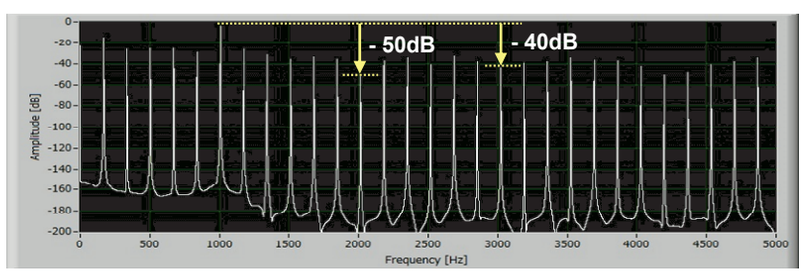
Fig. 18 Extracted waveform of sound signal (1 kHz, left channel). (a) Sobel filter, (b) LoG filter, (c) DoG filter.



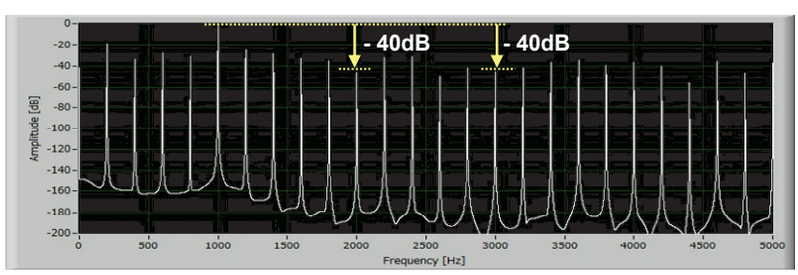
(a) Coventional record player.



(b) Sobel filter.



(c) LoG filter.



(d) DoG filter.

Fig. 19 Sound spectrum. (a) Conventional record player, (b) Sobel filter, (c) LoG filter, (d) DoG filter.

Table 3 Relative harmonic components of the reproduced sound (dB).

Reproduction Method		**Conventional Player**	**Sobel**	**LoG**	**DoG**
Hormonic Components	2nd.	-40	-30	-50	-40
	3rd.	-50	-30	-40	-40

6. Conclusions

This paper described the trial which reproduced the sound of analog records using digital image processing. First, we examined whether a groove geometry of stereo-record could be digitized based on the resolution of a flatbed scanner which was commercially available. And we confirmed that it was possible to reproduce the original audio information from the scanned image, although there were restrictions of dynamic range. Next, we tried three kinds of spatial filter processing (i.e. Sobel, LoG, and DoG filter), to the groove image captured by the scanner, and done comparison examination of those edge detection characteristics. As a result, we found that the LoG and the DoG filter showed relatively superior edge extraction characteristics. In addition, we tried reproduction of sound waveform from the image after each filter processing. We extracted waveform data of 1 kHz tone signal from the "Frequency Record", and created the WAVE files from these data. Then we evaluated the sound reproduction characteristics. As a result, we observed strong distortion components both in the time waveform and in the spectral waveform other than the original sound component of 1 kHz. The time waveform of the DoG filtered sound showed least high frequency noise components among the three filtering methods. In addition, the spectral waveform of the LoG filtered sound was most preferable because the harmonic distortion components were smallest. We are going to improve the accuracy of edge detection by using more sophisticated filter method.

It was found that sound recorded in Field-A (i.e. groove area in good condition in Fig. 5) was reproducible from the acquired image. It is necessary to enable reproduction of sound signal by a filtering process or the like for the sound recorded in Field-B (i.e. groove area in shadow in Fig. 5) in the future. And we would like to advance the study to allow for sound reproduction continuously by a unified algorithm from the sound track of the entire record. In addition, the scanner we used was for A4 (216×297mm^2). Therefore groove image of 30cm disk was obliged to image acquisition by dividing. About high precision groove connection of the images obtained with the division scan, we wish to consider it as a future subject.

References

(1) Stoddard, Robert E., "Optical turntable system with reflected spot position detection", U.S. Patent No. 4,870,631. 26 Sep. 1989.

(2) Stoddard, Robert E., and Stark R. N., "Dual beam optical turntable", U.S. Patent No. 4,972,344. 20 Nov. 1990.

(3) ELP Corporation, "Dear Analog Music Lovers", http://www.laserturntable.com/ (accessed on July 24, 2016).

(4) Uozumi, Jun, "Reproduction of Sound from a Monaural Disc by Means of Image Processing", *Bulletin of the Faculty of Engineering Hokkai-Gakuen University*, **35** (2008) pp.119-129 [in Japanese].

(5) Johnsen, Ottar, et al., "Detection of the Groove Position in Phonographic Images", *Proc. International Conference on Image Processing* (2007), pp.189-192.

(6) Springer, Ofer, "Digital Needle - A Virtual Gramophone", May 2002., http://www.cs.huji.ac.il/~springer/DigitalNeedle/ (accessed on July 24, 2016).

(7) Olsson, P., Ohlin, D., Olofsson, R., Ayrault C., and Vaerlien R., http://www.nyteknik.se/digitalisering/skannern-klarar-skivan-6437306 (accessed on July 24, 2016).

(8) Inoue, Toshiya (Eds.), *Record Player and Record*, (Rajio Gijyutsu Sha, Tokyo 1979), pp.45-180 [in Japanese].
(9) Sunier, John, *The story of stereo*: *1881*- (Gernsback Library, New York 1960), Chap. 5, p. 95.
(10) CONCERT HALL SM-972, Beethoven, Sonata No.14 in C Sharp Minor "Moonlight", ENTREMONT PHILIPPE.
(11) FREQUENCY VINYL DISC (LP) Ver. 2 (TYO-1003), Toyo Kasei Co. LTD.
(12) Young, I. T., Gerbrands, J. J., and Van Vliet, L. J., *Fundamentals of image processing* (1995), Chapter 9, ISBN 90-75691-01-7.
(13) Papari, G., and Petkov N., "Edge and line oriented contour detection: State of the art." *Image and Vision Computing* **29** (2011), pp.79-103.
(14) Marr, D., and Hildreth, E., "Theory of edge detection." *Proc. Royal Society of London. Series B*, **207** (1980), pp.187-217.
(15) Grigorieva, Ellina, *Methods of solving complex geometry problems*, (Birkhäuser, Basel 2013), p.132.