

タイトル	古蠟管からの画像工学的音声再生
著者	魚住，純；前田，尚範；吉田，拓馬；Uozumi, Jun； Maeda, Hisanori；Yoshida, Takuma
引用	工学研究：北海学園大学大学院工学研究科紀要(10)： 23-32
発行日	2010-09-30

# 古蠟管からの画像工学的音声再生

魚住 純\*・前田 尚範\*\*・吉田 拓馬\*\*\*

Image-engineering reproduction of sound from old wax cylinders

Jun Uozumi\*, Hisanori Maeda\*\* and Takuma Yoshida\*\*\*

## 1. はじめに

音を録音し再生する装置としての初めての蓄音機は、エジソンが発明したスズ箔蓄音機であり、その後、材質が蠟に変わり、蠟管式蓄音機が誕生した。蠟管は、音楽や口述の録音が比較的簡便にでき、複製も可能であることから広く普及し、音楽を聞く楽しみを一般人に解放した最初のレコードであった。また、音楽家らが自らの演奏を録音したり、言語学者が少数民族などの言語音声を録音するなど、研究目的にも多く使われ、貴重な音声資料としての蠟管も多数作成された。その中には現存するものも多いが、蠟管自体が貴重な文化財であることから蓄音機による再生が認められない場合や、割れや欠損などにより針による再生自体が不可能となっているものも多い。

そのような貴重な古蠟管から針を使わずに非接触で音声を再生する方法として、レーザービーム反射法が開発されている<sup>1-4)</sup>。これは、レーザービームを蠟管表面に照射し、それが反射する方向から音溝の傾斜を検知して音声情報を得るものである。この技術の開発の目的は、ポーランドの人類学者である B.ピウスツキが樺太アイヌの言葉を録音した蠟管から非接触に音声を再生することにあったが、その後、その他の蠟管の再生にも適用され、作曲家ヨハネス・ブラームス自身によるピアノ演奏の蠟管や、明治の演劇集団である川上音二郎一

座の琴の演奏などの歴史的音声の再生にも用いられている<sup>5)</sup>。この後者には、レーザービーム反射法を蠟管の金属型であるネガティブ用に改良した装置が使われている<sup>6)</sup>。

このように、レーザービーム反射法は蠟管からの実時間による音声再生法として有効な技術であるが、コヒーレントな光であるレーザー光を用いることから、光学的には明らかに粗面とみなされる蠟管表面からの反射によって生じるスペックルの影響が避けられず、再生音に広帯域のノイズが重畳することが大きな問題となっている。

一方、円盤レコードである SP 盤についても、損傷のあるレコードから貴重な歴史的音声を再生する目的から、光学的非接触法であるレーザー回折法が開発されている<sup>7)</sup>。この方式においても、スペックルがノイズの発生に加えて、レーザービームによる音溝のトラッキングにも支障を来すことが問題となっている。このため、レコードの音溝の顕微鏡拡大画像をデジタル的に処理することによって音声を再生する方法が開発され、部分的ながら良質の音声再生に成功している<sup>8)</sup>。本研究は、音溝画像のデジタル処理による音声再生法を古蠟管に適用することにより、蠟管の粗面状態やほこり等によって発生するスペックルノイズの影響を回避し、より良好な音質の音声を再生する方法を開発することを目的としている。本論文では、照射条件、照射光源、画像処理法等について検討を行っ

\* 北海学園大学大学院工学研究科電子情報工学専攻

Graduate School of Engineering (Electronics and Information Eng.), Hokkai-Gakuen University

\*\* 北海学園大学工学部電子情報工学科 (現在: 株式会社 iD)

Faculty of Engineering (Electronics and Information Eng.), Hokkai-Gakuen University (present: iD corporation)

\*\*\* 北海学園大学工学部電子情報工学科 (現在: 株式会社 NEC 情報システムズ)

Faculty of Engineering (Electronics and Information Eng.), Hokkai-Gakuen University (present: NEC Informattec Systems, Ltd.)

た結果, この方法による音声再生についての原理的な可能性を確認したので報告する.

## 2. 実験装置と音溝の撮像

### 2.1 資料と実験装置

エジソンタイプの蠟管は, 直径約 55 mm, 長さ約 105 mm の蠟製の管で, その表面に 254  $\mu\text{m}$  のピッチで最大 400 回転ほどの音溝が刻まれている. 回転数は初期のもので 144 rpm, 後期のもので 160 rpm であると言われるが, 必ずしも一定していなかったようである. 音声信号は音溝の深さの変化として記録されており, その深さは, 専門家による録音の場合に最大 50  $\mu\text{m}$  程度であり, 非専門家による録音ではそれよりも浅い場合が多かったとされている. 本報告の実験で用いた蠟管は, 製品として市販された音楽鑑賞用蠟管 (A) “Sun of my soul” および (B) “Just because she made dem Goo-goo Eyes” である. 始めに蠟管 (A) を用いて実験を開始したが, 途中で破損したため, 蠟管 (B) に切り替えて実験を継続した. 図 1 (a) に蠟管 (B) の写真を示す. 同図 (b) の拡大写真からは, 音溝が深さ方向に振動していること, および蠟管表面が部分的に白く変色していることが確認できる.

この音溝の深さ方向の変化を画像として捉えるため, 蠟管表面に斜め方向から光を照射し, 音溝形状によるその陰影を単眼ズーム式顕微鏡 (シグマ光機, MXZ-2) により拡大して, カメラにより撮像する. カメラには, 画素数 1280 $\times$ 1024 の CMOS 型白黒カメラ (アートレイ, ARTCAM-130MI-BW) を使用した. z 軸パルスステージ (シグマ光機, SPSG26-200) にマウントした回転パルスステージ (シグマ光機, SPSG-60YAW-OB) の上に蠟管を中心軸が垂直になるように置き, 回転ステージにより蠟管を一定角度ずつ回転させながら撮像を行い, 1 周の回転の後, z 軸方向に一定距離移動させて, 2 周目の撮像を行うという方法により, 蠟管全体にわたる音溝画像の取得を行った. 2 つのパルスステージは, 2 軸ステージコントローラ (シグマ光機, SHOT-102) を USB 経由の GP-IB (National Instruments, GPIB-USB-HS) を介して PC から制御することにより駆動した. 用いた光学系の写真を図 2 (a) に示す. 図 2 (b) は, パルスステージ上の蠟管と光照射部を示

したものである.

照射光の光源には, 光ガイド (中央精機, LGG-500WF) により導光したハロゲンランプ (中央精機, SPH-100N) の光に加え, 波長 532 nm, 出力 1-3 mW の小型緑色固体レーザー (B&W Tek, BWE-3E), および中心波長 470 nm の高輝度青色 LED スポットライト (Edmund Optics, 63336-L) を用いた. このうち, レーザは, ビーム径が 1 mm $\phi$  と小さいため, 6 倍のビームエキスパンダ (Edmund Optics, 28272-L) を用いて拡大した. 蠟管表面への光の照射に用いたこれら 3 つの光源を図 3 に示す.

撮影した画像は, USB を介して PC に取り込み, MATLAB を用いて画像処理および信号処理を行った. PC には, プロセッサ Core2Duo E6400 (2.13 GHz), メインメモリ 1.93 GB に Windows XP Professional を搭載したものをを用いた. MATLAB は, Image Processing Toolbox (IPT) および Instrument Control Toolbox (ICT) を併用し, GP-IB の制御には ICT の関数を用いた.

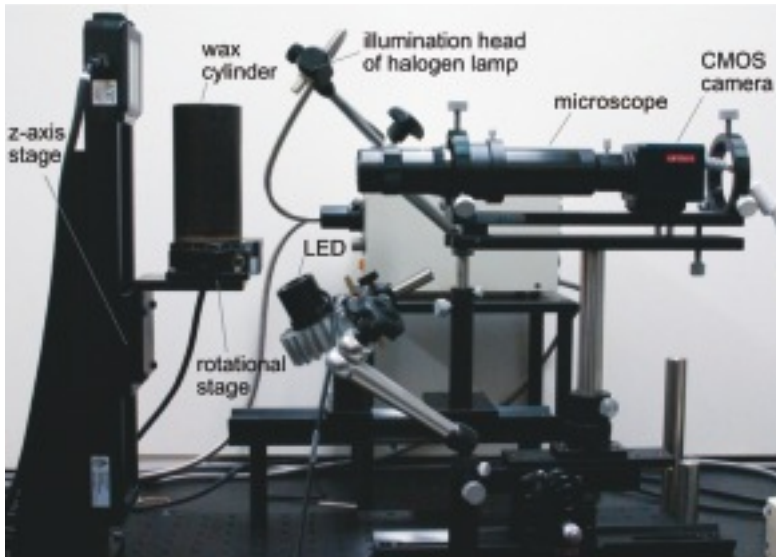


(a)



(b)

図 1 (a) 蠟管 (B) と (b) その部分拡大



(a) 光学系と蠟管駆動部



(b) パルスステージと照射部

図2 蠟管の音声再生のための光学系

(a) ハロゲンランプ，光ガイド  
および照射ヘッド(b) ビームエキスパンダを装着  
した小型緑色固体レーザー(c) 高輝度青色LED  
スポットライト

図3 照射に用いた3つの光源

## 2.2 音溝の撮像

音溝の起伏を明確に反映する陰影を生成するため、蠟管表面に外接する円筒面の法線に対して下側の斜め方向から光を照射する。入射角  $\theta$  を  $20^\circ$  から  $60^\circ$  まで変えながらハロゲンランプの光を照射した場合の音溝画像の変化を図4(a)–(e)に示す。この照射方向では、主として音溝の上側壁面に光が当たり、下側壁面が暗部となる。このため、明部の上側と暗部の下側の境界線が、音溝の上縁の形状を反映していると考えられる。

蠟管の音溝は、先端が球形の記録針が上下に振動することによって刻まれるため、音溝が深いほど幅が広がっている。したがって、このような照射法によって音溝の縁の形状を捉えることにより、それを深さに読み替えて音声再生することが可能であると推測される。深さと幅の関数関係は

一般に線形である保証はないが、その変化が比較的小振幅であること、および蠟管の録音・再生特性自体がいわゆる Hi-Fi ではないことから、その非線形性は大きな問題とはならないと考え、この照射方式を用いることとした。

図4に見られるように、入射角が小さいと、音溝の底部近くまで光が届き、暗部の幅が狭くなるとともに、音溝の下側壁面の近くまで輝度が高くなり、一方、入射角が大きいと、音溝の上側壁面に光が当たらない部分が生じる確率が大きくなるのが分かる。本研究では、 $\theta \approx 60^\circ$  が適当であると判断し、以後この入射角を用いることとした。

円盤レコードからの音声再生においては、ハロゲンランプを使用することで良好な再生音声を得られていることから、本実験においても始めにハロゲンランプの使用を試みた。しかし、蠟管が円筒形状であるため、光が照射ヘッドから広角度に

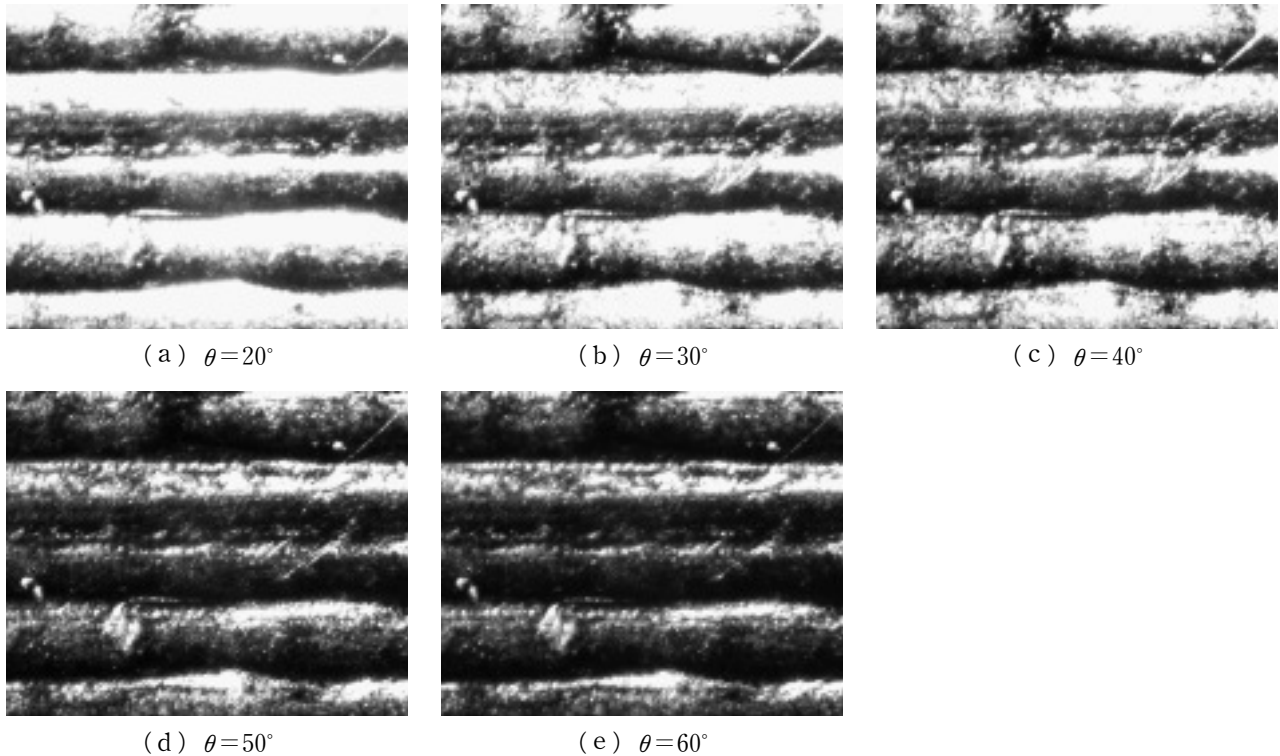


図 4 入射角  $\theta$  による音溝画像の変化

拡散するハロゲンランプでは、同一音溝部位を隣接する画像に重複撮影した場合に、その照射条件が大きく変化し、抽出した音声情報の接合に不都合が生じる可能性が考えられる。また、蠟管の音溝は、円盤レコードと比較して浅いため、発散角が大きい光ビームでは陰影の明確な境界が生じにくいことも想定される。このため、ほぼ平行光と見なせるレーザー光、および発散角の小さい高輝度 LED スポットライトを用いて撮像を行った。

図 5 (a) および (b) にハロゲンランプ光の照射ヘッドからの拡散状態と実際に蠟管に照射した際の照射光の広がりを示す。同図 (c) および (d) は高輝度 LED スポットライトについて示したものである。また、同図 (e) は、小型固体レーザー光を照射した状態を示している。図 5 (a) および (c) は、ほぼ光軸をとる平面内に白紙をおいて撮影し、光が伝搬する様子を擬似的に示したものである。この場合、光は紙面で散乱されながら進むため、光は伝搬するに従って実際よりも強く減衰して見えていることに注意する必要がある。これらの写真から、光ガイドから出射するハロゲンランプの光の発散角は大きく、蠟管を広範囲に照射しているのに対し、高輝度 LED スポットライトの光は比較的平行性が高く、蠟管への照射もハロゲンランプの場合ほど広がらないことがわかる。一

方、小型固体レーザー光は、6 倍のビームエキスパンダを使用してもビーム幅は数 mm と小さく、レーザー光故に平行性も高いことから、蠟管への照射領域が最も狭いことが分かる。

以上のことから、蠟管全体の音声再生には、レーザーと LED を光源として用いることとした。1 週の撮影枚数や画像 1 フレームあたりの音溝数は、設定値を変えながら実験をすすめたが、最終的な音声再生においては、顕微鏡の倍率を 3 倍とし、1 画像中の音溝数 7 本、1 周あたりの画像数 100 枚、総撮影周数 51 周、総計 5100 枚の画像として記録した。なお、音溝から回復した音声信号を回転方向に隣接する画像からの音声信号と接続する際に接続点を確認できるようにするため、隣接画像間で音溝の同一部分が重複して含まれるように撮影フレームを設定した。また、取得した画像は、画質を保持しつつデータサイズを抑制するため、高精細の JPEG 画像として保存し、後の画像処理に供した。

### 3. 画像処理

#### 3.1 スペックルの低減

レーザー光照射により撮像した音溝画像とその部

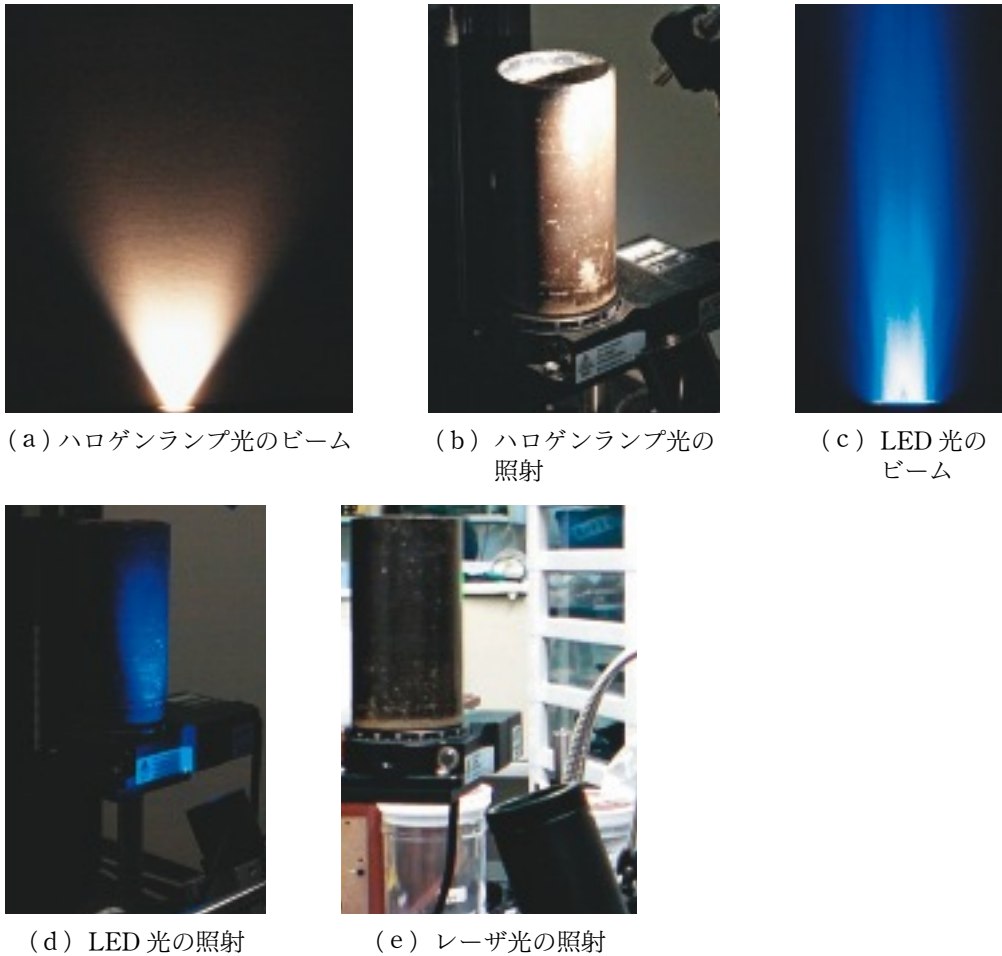


図5 照射光の広がりと蠟管への照射状態

分拡大画像の例を図6(a)および(b)に示す。また、図7(a)および(b)は、ほぼ同一の場所をLED照射により撮像した結果である。図6が示すように、コヒーレント光であるレーザー光を照射した場合には、音溝の陰影画像が蠟管の粗面特性に起因するランダムな干渉縞であるスペックルによって強く変調されていることが分かる。一方、LED光照射による画像には、表面の微細な凹凸による斑点状の明暗は認められるものの、光のコヒーレンスが極めて低いことから、スペックルは発生していない。

レーザー光照射によって生じるスペックルは、いわゆる十分に発達したスペックルであると考えられる。その場合、強度確率密度関数は零強度を最大とする負指数分布に従う<sup>9)</sup>。このため、背景となる低強度領域の面積が広く、図6の画像をそのまま2値化すると、黒い背景に無数の白い孤立スペックル粒が浮かぶ画像となることから、2値化の前処理としてスペックルの低減が必要となる。これに対し、LED照射画像に見られる斑点模様

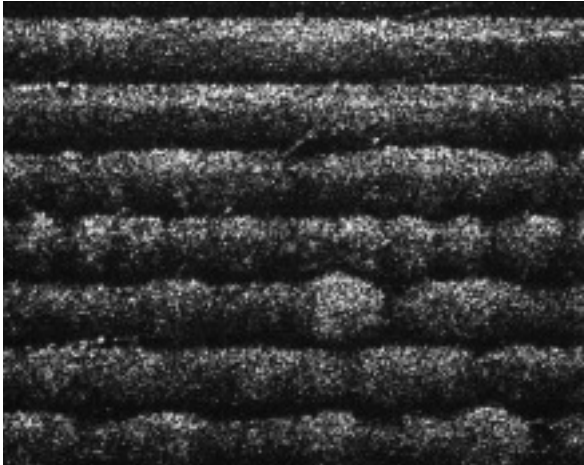
は、2値化の閾値を適切に設定すれば大きな問題とはならないと考えられる。

スペックルの低減は、画像のフーリエスペクトルにガウス型ローパスフィルタを適用する空間周波数フィルタ処理により行った。ガウス関数の $x$ 幅は $x$ 方向の基本空間周波数の30倍とし、 $y$ 幅は画像の縦と横の画素数比を考慮して、 $x$ 幅と同一の空間周波数となるように設定した。図6とは別の部位の音溝画像である図8(a)に対してスペックル低減を行った結果を図8(b)に示す。スペックル粒間の低強度部分が平滑化により消失していることが分かる。

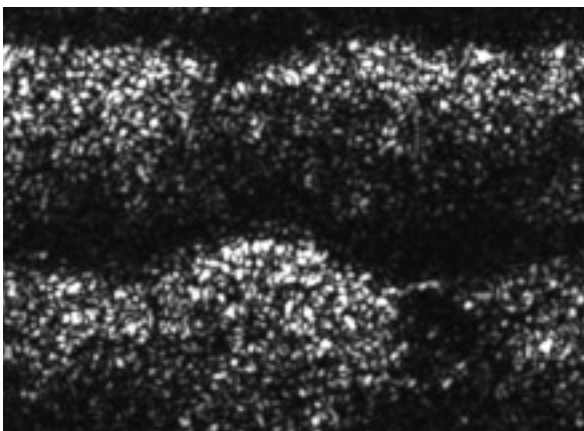
スペックルの低減以降の処理は、LED光照射による画像の場合とほぼ同一であるため、以下ではレーザー光照射による画像を使って処理の説明を行う。

### 3.2 2値化と音溝形状の整形

音溝の上縁を表す明部と下縁を表す暗部との境



(a) 全体



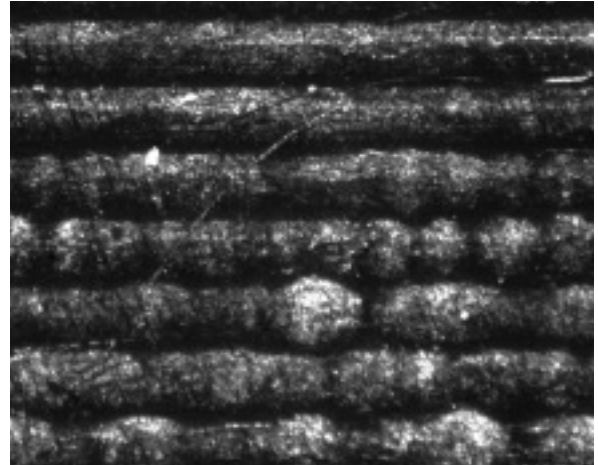
(b) 部分拡大

図6 レーザ光照射による音溝画像

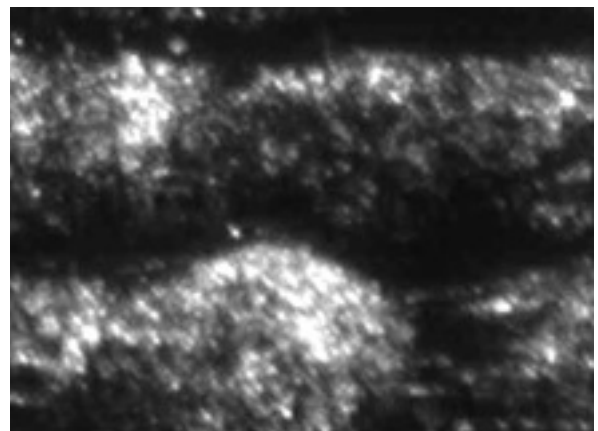
界線を明確にするため、2値化を行う。蠟管の画像は、部位によって蠟の地の色が異なる場合があることから、閾値を画像毎に個別に定めることとし、各画像の平均濃度値に0.8-0.9程度の係数をかけた値を閾値として設定した。平均濃度値の0.9倍を閾値として図8(b)を2値化した結果を図9に示す。

2値化された画像には、黒地中の白い斑点、白地中の黒い斑点、突起、欠損など、本来の音溝形状に対応しない図形成分や形状が多く含まれている。これらの不要な要素を可能な限り除去するため、小領域図形成分の除去、膨張・収縮処理、およびその後の形状調整を行った。

2値化後の音溝画像から、一定面積以下の黒および白の小領域図形成分を除去した結果を図10(a)に示す。この処理には、MATLABのIPT関数である**bwareaopen**を使用した。この処理により、小領域の白および黒の図形成分は概ね取り除かれているが、少数ながらより面積の大きな孤



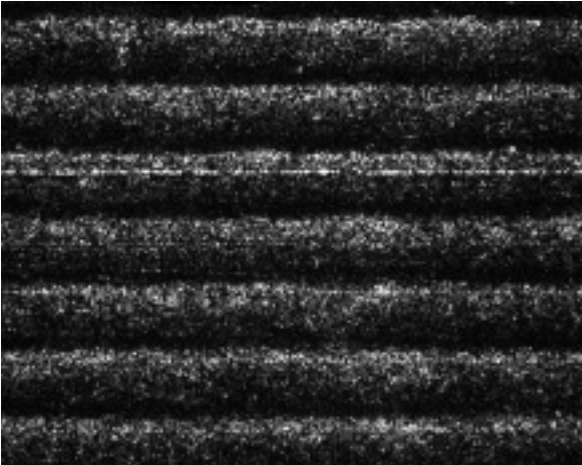
(a) 全体



(b) 部分拡大

図7 LED光照射による音溝画像

立図形成分がまだ残っており、白の図形成分から黒地に出ている突起や白地の欠損部分も、正しい音声情報には対応しないと考えられる。このため、膨張と収縮を組み合わせることで、それらの除去を行った。音溝画像は、 $x$ 方向が時間軸に対応し、 $y$ 方向が音声の振幅軸に相当するという意味で等方的ではない。したがって、これら2つの方向に対して異なる膨張・収縮画素数を指定して1次元の膨張・収縮を繰り返し行った。すなわち、音声信号への影響を抑えつつ、突起や欠損による音声信号の異常値発生を抑制するため、 $y$ 方向には $x$ 方向の半分の画素数を設定した。使用した関数は、IPT関数の**imerode**および**imdilate**である。図10(a)の画像にこの処理を施した結果が同図(b)である。2つの方向に対して別個に処理を行っているため、処理後の画像は図形成分の輪郭が角張った形状になっている。この輪郭に丸みを持たせるため、関数**medfilt2**を用いてメディアンフィルタ処理を行った結果を図10(c)



(a) 原画像

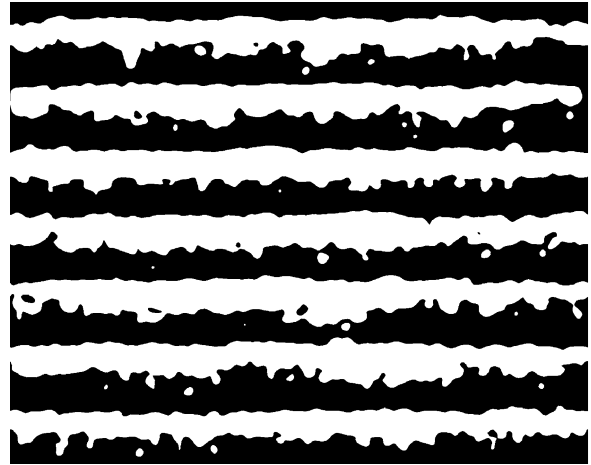
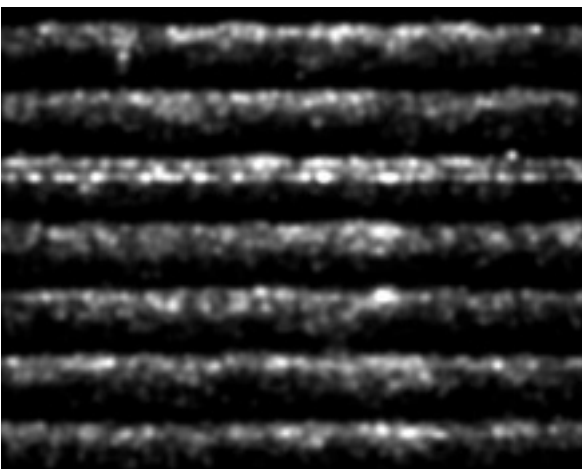


図9 図8(b)を2値化した画像



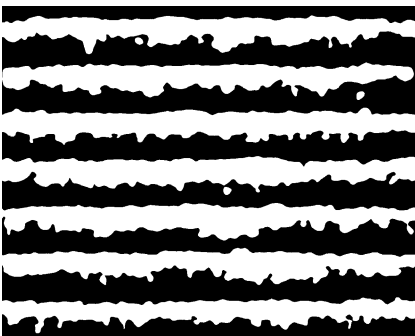
(b) スペックル低減後の画像

図8 空間周波数フィルタによるスペックル低減

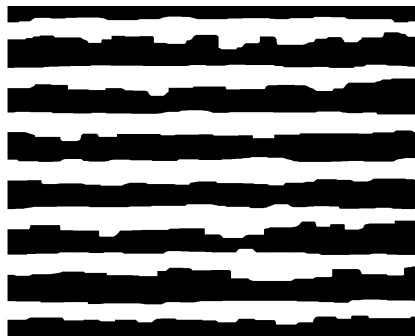
に示す。なお、この処理は白図形の左右端も丸める特性があるため、その部分を切り捨ててその後の処理を行った。

### 3.3 音声信号の抽出

以上の処理により、1つの画像中に含まれる7本の音溝の縁の形状を反映した境界線が生成されている。この画像から各音声信号を抽出するためには、7本の境界線を含む帯状領域を切り出す必要がある。音溝の撮像に当たっては、各画像中の音溝の位置がほぼ一定となるように、z軸パルスステージの移動量を設定している。このため、帯状領域の切り出し位置は全ての画像に共通とし、画像の行成分を7等分して各成分の上半分を切り出し領域とした。この方式により図10(c)の画像から切り出した7本の帯状領域を図11に示す。この各領域内の白と黒の境界線の高さが音声信号に相当している。その抽出には、各画像を表す配列について、列方向に和をとる方法を用いた。その結果得られる1次元配列が一つの帯状領域に対応する音声信号のセグメントとなる。



(a) 小領域図形成分除去後



(b) 膨張・収縮処理後



(c) メディアンフィルタリング後

図10 2値化音溝画像の整形処理



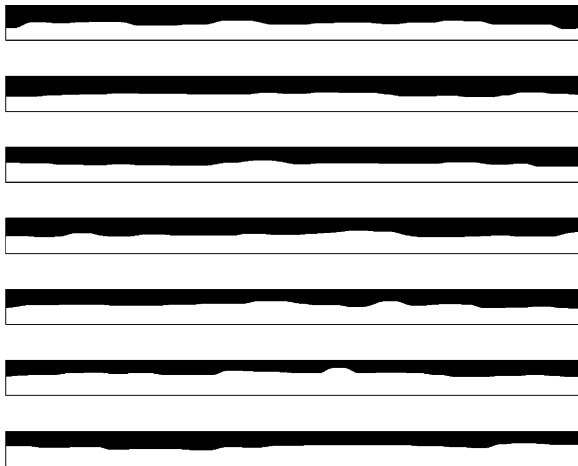


図 11 音溝の上縁を含む帯状領域の切り出し

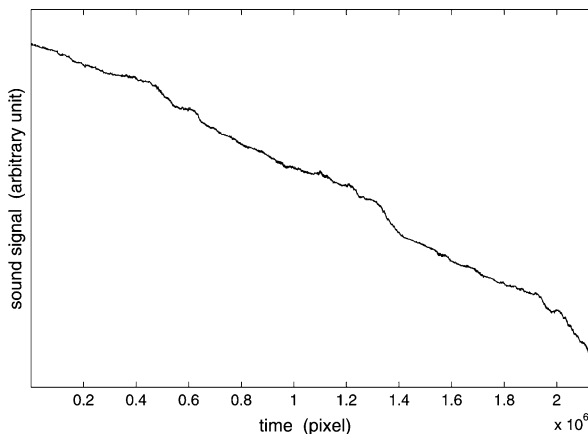
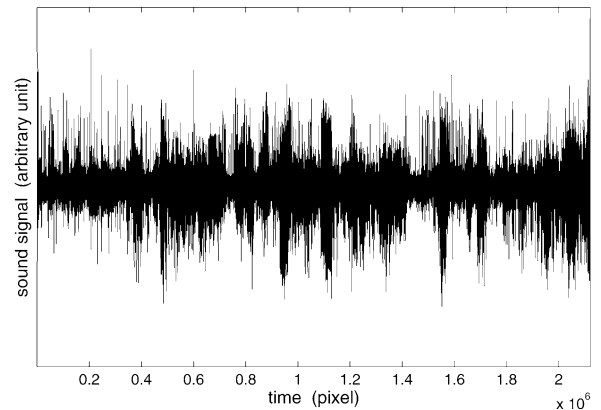


図 12 音声信号セグメントの接続により合成した信号

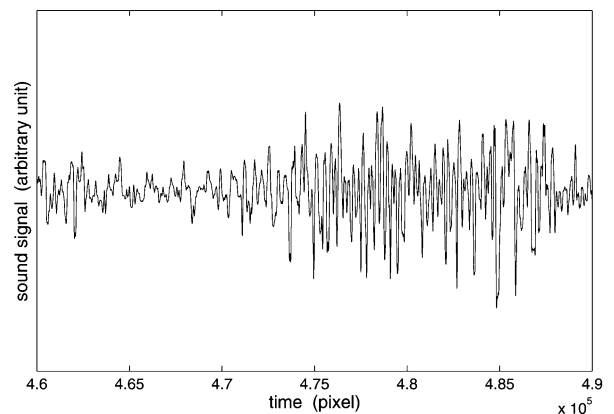
### 3.4 音声信号セグメントの接続

この音声信号の各セグメントは、隣接する画像から抽出された対応する各位置のセグメントと接続し、最終的には全画像から抽出した全てのセグメントを接続して一つの音声信号を再構成する。前述のように、各画像は撮影時に隣接画像と一部重複するようにフレームを設定しているが、その重複距離を一定の値としていることから、その距離に相当するデータを各セグメントから削除して、順次接続する処理を行った。ただし、カメラや回転ステージが正しく水平に設置されていない、あるいは蠟管の端面が回転軸に対して垂直になっていないなどの原因から、接続すべき隣接セグメント間に高低差が生じる可能性がある。このため、接続に際して、接続点での信号値の差に相当するバイアスを接続するセグメントに加えて接続を行った。

PC のメモリサイズの制限から、実際には全て



(a) 全体の波形



(b) 一部分を拡大した波形

図 13 第 22-24 週の画像から再構成した音声信号

のセグメントを一度に結合することが困難であったため、蠟管 1 本分を撮影した 51 週のうち、50 周 5000 画像を 3 周 300 画像ずつのブロックに分割し、各ブロック内の音声信号セグメントを上述の方法により接続した。51 周目の画像には、音溝が 2 本しか含まれていなかったため、音声再生処理には含めなかった。

例として、第 22-24 週の画像のセグメントを接続した結果を図 12 に示す。この信号には、全体に右下がりのトレンドが含まれている。これは、上述のバイアス調整を行ったことにより発生したものと思われる。このトレンドを除去し、さらに蠟管に記録されているとされている約 100 Hz-3.5 kHz 帯域のみを残す帯域通過フィルタ処理を、FFT によるフーリエ空間において行った。その結果得られた音声信号を図 13(a) に示す。同図 (b) は、その一部を横に拡大して示したものである。

このようにして得られた画像 3 周ごとの音声信号を、データ形式をリニア PCM としたまま、Microsoft Windows の標準音声ファイルである

WAV ファイルに保存した。その際、蠟管の回転数を 144 rpm と仮定して、離散的音声データのサンプリング周波数を算出した。それらを波形編集ソフトウェアを用いて一つのファイルに結合することにより、最終的に蠟管 1 本分の音声の一つの音声ファイルに再構成した。

#### 4. 結果と考察

再構成した音声の長さは約 2 分 17 秒であり、これは標準的な蠟管の録音時間の範囲内である。録音内容である “Just because she made dem Goo-goo Eyes” は、蠟管から蓄音機により再生して MP3 形式に変換したファイルがインターネット上に著作権フリーのファイルとして公開されており、この曲が楽器演奏のないボーカルだけの楽曲であることが確認されている。本研究により再生した音声は、雑音のレベルが高く、歌詞を聞きとるのはかなり困難であるものの、そのメロディがほぼ正確に再生されており、再生時間も公開されている音声とほぼ同じである。その意味で、斜め光照射による音溝の陰影画像から画像処理により音声を再生することが、原理的には可能であることが確認されたと言える。

再生音の品質はまだ低く、今後の改良が必要である。本研究では、再生音声の生成を第一目標に実験を行い、画像処理等におけるパラメータは適当と思われる暫定値を使用した。今後、一連の処理の中で用いている方法やパラメータを精査し、最適化することにより、音質向上は可能であると考えられる。

なお、本実験では、高輝度 LED 光照射による画像とレーザー光照射による画像に対して、ほぼ同一のパラメータを用いて画像処理を行い、音声ファイルを生成した。その結果、両者の再生音には大きな違いは感じられなかったが、今後音質の向上に伴って違いが出てくる可能性があり、引き続き両者の比較検討が必要である。

#### 5. おわりに

蠟管の表面に斜め方向から光を照射して生成した音溝の陰影画像をデジタルカメラにより PC に取り込み、画像処理を施すことにより、その音声を非接触で再生する方法を提案し、それが原理的に可能であることを確認した。

同様の方法は、すでに SP 盤レコードからの音声再生に適用されており、良質な音声再生が可能であることが示されている。蠟管と円盤レコードでは、音溝への音声の記録方法が異なることから、蠟管に本方法を適用することは円盤レコードの場合よりも難しいと考えられるが、その可能性が示されたことの意義は大きい。

蠟管の非接触音声再生法としては、レーザー光を用いたレーザービーム反射法がすでに開発されている。レーザービーム反射法が実時間再生法である反面、スペckルに起因するノイズレベルの高さが問題となっているのに対し、画像処理による本方法は、撮像と画像処理に多大の時間を必要とするものの、画像処理の改良によってレーザービーム反射法を上回る低ノイズ、高音質の実現できる可能性を秘めている。したがって、今後ノイズの低減と音質の改善に向けての光照射方式および画像処理の最適化は必須の課題である。

本研究は科学研究費補助金(基盤研究(A))「蠟管を中心とした初期録音資料の音源保存・音声復元・内容分析に関する横断的研究」の支援のもとで行われた。

#### 【参考文献】

- 1) 伊福部達, 朝倉利光, 川嶋稔夫: ピウスツキ録音蠟管レコードとその再生方法, 日本音響学会誌, Vol. 40, No. 3, pp. 167-174, 1984.
- 2) 朝倉利光, 岩井俊昭, 川嶋稔夫: レーザー光を用いたピウスツキ録音蠟管からの光学的音声再生, O plus E, No. 66, pp. 64-72, 1985.
- 3) T. Iwai, T. Asakura, T. Ifukube and T. Kawashima: Reproduction of sound from old wax phonograph cylinders using the laser-beam reflection method, Appl. Opt., Vol. 25, No. 5, pp. 597-604, 1986.
- 4) J. Uozumi and T. Asakura: Optical methods for reproducing sounds from old phonograph records, *International Trends in Optics and Photonics ICO IV* (Springer, Berlin, 1999) pp. 409-425.
- 5) 魚住 純, 牛坂 健, 伊福部達: 古ろう管からの名演の復活, 日本音響学会誌, Vol. 54, No. 6, pp. 443-449, 1998.
- 6) 牛坂 健, 魚住 純, 朝倉利光: ネガティブ蓄音管の光学的再生装置, 北海学園大学工学部研究報告, No. 26, pp. 383-400, 1999.
- 7) J. Uozumi and T. Asakura: Reproduction of sound from old disks by the laser diffraction method, Appl. Opt., Vol. 27, No. 13, pp. 2671-1986, 1988.

- 8) 魚住 純：画像処理によるモノラル円盤レコードからの音声再生，北海学園大学工学部研究報告，No. 35, pp. 119-129, 2008.
- 9) J. W. Goodman: Statistical properties of laser speckle pattern, in *Laser Speckle and Related Phenomena, Second Enlarged Edition*, ed. J. C. Dainty (Springer, Berlin, 1984) pp. 9-76.