

超音波映像の AR 表示を用いたグロウル歌唱学習支援

田口諒† 宮下芳明†

概要: グロウルとは、主にメタル音楽で用いられる激しく喉を震わせる歌唱方法である。グロウルは通常の歌声と大きく異なる歌唱方法であるため、初心者にとって練習時の歌唱を自己評価することが難しい。本研究では、喉頭と横隔膜の超音波映像を用いることでグロウルの学習を支援する。超音波映像装置は身体的な負担が少なく、リアルタイムに体内を観察できる。また、超音波映像と身体的位置の対応付けを分かりやすくするために AR 表示することとした。

1. はじめに

本研究では、グロウルの効率的な上達を目的とし、超音波映像を用いて初心者の学習を支援する。超音波映像装置は身体的な負担が少なく、リアルタイムに体内を観察できる[1,2,3]。また、超音波映像と身体的位置の対応を分かりやすくするために AR 表示することとした。喉頭と横隔膜の超音波映像の AR 表示を図 1, 2 に示す。超音波映像の AR 表示を左側に提示し、右上に元の超音波映像、右下に AR 表示の透明度を調整するスライダーを提示する。システムを利用した練習風景を図 3 に示す。外付けカメラに AR マーカが映るように調整し、提示画面を見ながらグロウルの練習を行う。

グロウルとは、主にメタル音楽で用いられる激しく喉を震わせる歌唱方法である。日本語ではデスボイスと呼ばれることが多い。マキシマム ザ ホルモンや Crystal Lake などのバンドは、曲中にグロウル歌唱を取り入れている。グロウル歌唱時は声帯以外の部位が同時に振動することや(図 4)、長期間の使用でも喉に悪影響のないことが分かっている[4,5]。また、グロウルは通常音声よりも声帯を閉じる力が弱く、息を吐き出すスピードが速いことも分かっている[6]。

教則本ではグロウルの歌唱方法について、犬のうなり声を真似る、喉を締めすぎず傷めないように発声するなど説明されている[7,8]。また、声に徐々にエッジを足していく感覚で、喉に負担をかけずに練習を行う方法も存在する[9]。グロウルを習得することで、歌唱の幅が大きく広がり、より多様なパフォーマンスが可能になる。しかし、グロウルは特殊な歌唱方法であるため教わる機会が少ない。そのため多くの方はグロウルを独学で学ばなければならない現状がある。独学の場合、ウェブの解説動画や教則本、グロウル歌唱を含む楽曲などから知識を得ることになる。

グロウル初心者 6 名に独学でグロウルを練習させる調査を行った。その結果、自身による歌唱が正しいグロウルで

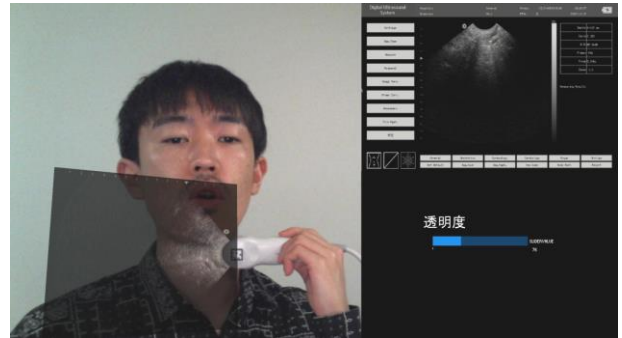


図 1. 喉頭の超音波映像の AR 表示。

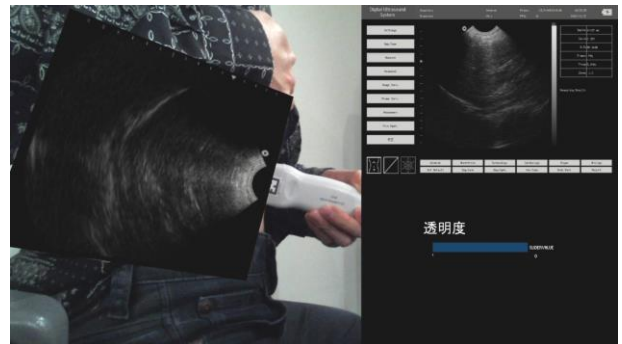


図 2. 横隔膜の超音波映像の AR 表示。

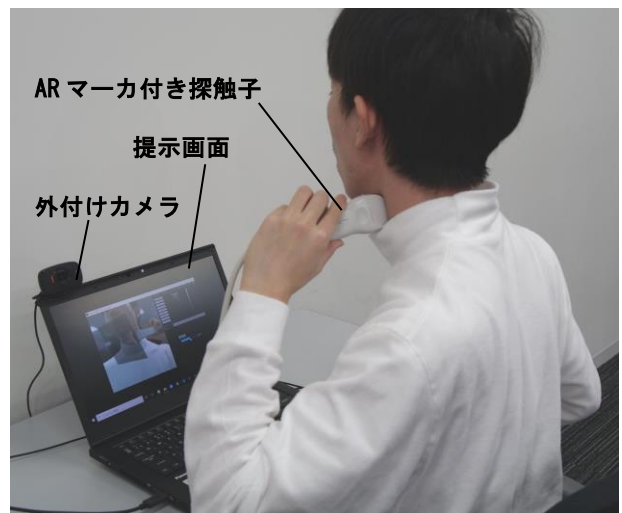


図 3. システムを使った練習風景。

あるかどうかを、歌唱しながら聴いて判断することが難しいという意見が得られた。また、グロウルはどこを直せばよいか分からないので非常に難しいという意見も得られた。自己評価が誤っている場合、グロウルが思うように上達しない原因になりうる。また、正しく発声できているつもりでも実際は喉を傷める発声になっている可能性がある。

カラオケには、現在の歌唱に対する評価が画面上に提示される機能がある。リアルタイムかつ視覚的なフィードバックは歌唱の練習において重要視されている[10,11,12]。本研究では、グロウル歌唱時の超音波映像から得られる情報について調査・整理し、超音波映像のAR表示を利用したグロウル学習方法について考察する。

2. 超音波映像から得られる情報

超音波映像は主に医療現場での検査方法として用いられ、エコー検査と呼ばれることが多い[1,2,3]。探触子を皮膚に接触させることで、リアルタイムに体内の超音波映像を観察できる。超音波とは人間の可聴域(約20Hz~20kHz)を超える高い周波数の音波である。この超音波の反射時間を利用して対象までの距離を推定し映像化する。超音波映像装置は超音波の直進性や減衰、屈折、反射など様々な要素を考慮して、設定を調整する必要がある。

超音波映像は頸部の広い領域をリアルタイムに観察できる[2]。ただ、下顎骨が下がった場合上頸部の情報を得ることが難しいため、観察時の姿勢の工夫や頸部の観察に適した形状の探触子を使用する必要がある。Naokiらが口腔の超音波画像を利用して無発声音を検出するシステムを提案している[13]。顎の下部に超音波映像装置の探触子を当てることで、舌や舌骨などの口腔周辺の情報を取得している。探触子を肋骨に当てることで、横隔膜を観察できる[3]。横隔膜の下部には肝臓を観察することができる。グロウル歌唱の質に喉頭や横隔膜の状態が関係しているため[4,6]、それらの組織を観察することができる超音波映像を支援情報として利用する。

3. 超音波映像におけるグロウルの特徴の調査

3.1 概要

グロウル発声時の喉頭と横隔膜の超音波映像を調査した。喉頭は声の生成に関係する器官であり、横隔膜は呼吸に関係する筋肉である。観察には超音波映像装置(CMS600P2)とコンベックス走査型探触子(Micro-Convex Cardiac Probe)を利用した(図5)。探触子を当てる位置を図6に示す。

3.2 超音波映像装置

送受信された超音波は超音波映像装置(CMS600P2)によって計算され、映像としてモニターに映し出される。本研究で用いる探触子はコンベックス走査型探触子(Micro-Convex Cardiac Probe)で、視野が広く、深くまで観察できることが特徴である。対象を超音波でなぞり情報

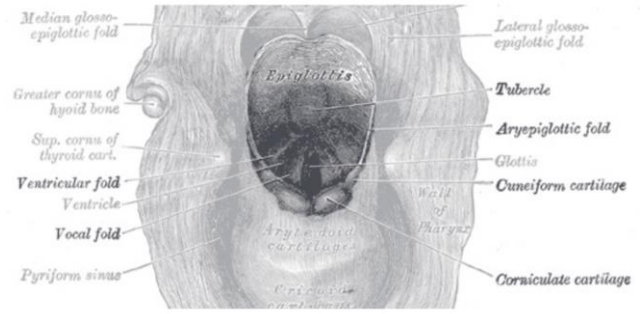


図4. メタルで用いられるような激しいグロウルは、歌唱時に濃い文字で示された部位が振動する(〔4〕より引用)。



図5. 超音波映像装置(CMS600P2)とコンベックス走査型探触子(Micro-Convex Cardiac Probe)。

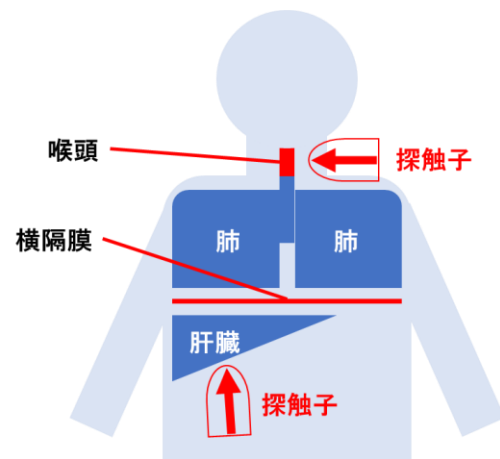


図6. 喉頭と横隔膜周辺の概略図と探触子の当て方。

を得ることを走査と呼び、超音波映像装置では様々な走査モードが存在する。本研究で用いる走査モードはリアルタイムで断層を表示できるBモード(断層モード)である。Bモードにおける各種設定について表1に示す。周波数やフォーカス位置、ゲインなどの設定によって、より観察対象を見やすく調整できる。

超音波映像の横方向解像度は近い対象物で2mm、遠い対象物で4mmである。また、軸方向解像度は2mmである。

表 1. B モードにおける調整 (参考文献[1]より抜粋引用).

設定項目	操作	画質
周波数	上げる	分解能：上がる 深部感度：下がる
	下げる	分解能：下がる 深部感度：上がる
フォーカス	目的の位置 に合わせる	焦点が合った 画像になる
ゲイン	上げる	全体が明るくなる
	下げる	全体が暗くなる

接触面から 3mm 以内は仕様上、映像化できない。

焦点距離を変えた複数の超音波を送信することで分解能を向上させる方法は多段フォーカスと呼ばれる。また、フレーム間の平均をとることで、物体の概形を見やすくする方法も存在する。しかし、これらの方法はフレームレートが低下する原因となるため動く物体の観察に適さない。

3.3 喉頭周辺の観察

著者 1 名を含めたグロウル歌唱経験者 4 名 (20 代, 男女各 2 名) を対象に喉頭周辺の超音波映像を調査した。喉頭は声帯や仮声帯などを含む、声の生成にかかわる器官である。体表からは喉仏として触れることができ、嚥下時には一時的に上部に移動する。

図 6 の喉頭位置に探触子を当てたときの超音波映像を図 7 に示す。グロウル歌唱時はクリーン歌唱 (通常の歌唱) 時に比べて、図 7 の赤丸で示す部分が激しく動くことが分かった。これは、グロウル歌唱時に声帯以外の組織が振動することによって引き起こされている可能性がある[4]。この現象は 4 名すべてのグロウル歌唱経験者で確認された。

男性 2 名は喉頭を映し出すことが容易であったが、女性 2 名は喉仏が出ていないため、位置を探すことに時間がかかった。基本的に嚥下を行うことで、喉仏が上下するため、それを目印に喉頭を探すと見つけやすい。また、超音波映像と身体の位置の対応の把握が難しく、見る位置があっているのか質問する参加者がいた。超音波映像と身体の位置の対応を把握できる提示方法を検討する必要がある。

3.4 横隔膜周辺の観察

著者 1 名を対象に横隔膜の超音波映像を調査した。横隔膜は胸腔と腹腔の間に位置する呼吸にかかわる板状の筋肉である。横隔膜が下がると肺が膨らみ、吸気される。横隔膜が上がると肺が収縮し、息が吐きだされる。横隔膜は肝臓の頭側に位置するため、肝臓近くの肋骨に探触子を当てることで観察することができる (図 6) [3]。

横隔膜の超音波映像を図 8 に示す。グロウル歌唱時はクリーン歌唱時に比べて、図 8 左図から図 8 右図の状態に移り変わるまでが早かった。グロウルの方が、息を吐くスピードが速いためだと考えられる[6]。呼吸の有無だけでなく、

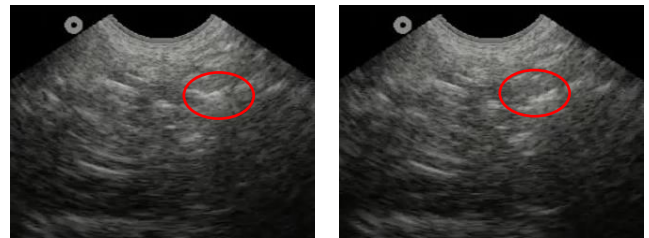


図 7. 喉頭の超音波映像。グロウル歌唱時に赤丸で囲った部分が激しく動く。

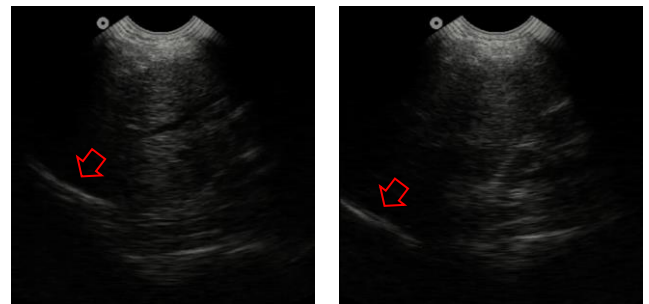


図 8. 横隔膜の超音波映像。息を吐くにつれて、左図から右図の状態に遷移する。クリーン歌唱・グロウル歌唱時同様。

吸い込んだ息の量や息を吐き出すスピードの相対的な違いをセルフチェックできるという点で、グロウル歌唱の学習に利用できると考えている。

4. 提案手法

4.1 概要

調査の結果、喉頭と横隔膜の超音波映像においてグロウルの特徴が観察できることが分かった。一方で、超音波映像と身体の位置の対応の把握が難しいことが分かった。観察対象の位置に超音波映像を重畳表示することで、位置の把握を補助しながらグロウルの学習を支援する (図 1, 2)。発声練習時の喉頭と横隔膜を見本映像と見比べて、グロウルの特徴が表れているかを確認する。これにより声の生成と息の使い方の上達をセルフチェックできる。

4.2 システム

プログラムを実行すると図 1, 2 のような画面が提示される。超音波映像装置 (CMS600P2) のコンベックス走査型探触子 (Micro-Convex Cardiac Probe) による入力を利用している。キャプチャデバイス (GV-HUVC) を用いて超音波映像装置の映像出力を PC (Let's Note CF-LV9KDNQR) に入力している。AR マーカを用いて探触子の位置と角度を認識し、超音波映像を AR 表示している。AR マーカまでの距離に応じて、超音波映像の AR 表示が自動的に拡大・縮小する。AR 表示の透明度を変更することで、背景のカメラ映像の見え方を調節することができる。透明度は 0~255 の数値がスライドバー右下に提示されており、255 で完全に透明になる。外付け Web カメラの位置を変えるこ

とで、図2のような角度からも観察することができる。用意された見本の動画ファイルは必要に応じて再生する。

4.3 利用方法

グロウルは叫んでいるように聞こえるため、叫ぶことから発声練習を始める初心者もいるが、喉を傷める危険性があるため、本研究では段階的な発声練習を想定している。発声しながらシステムによる提示画面を確認することで、リアルタイムに自身の体内の様子を確認できる(図3)。グロウルの見本を提示画面のとなりと並べて再生し、見本に近づけていくイメージで練習を行うことができる。さらに、喉頭や横隔膜に絞ったチェックができるため、何ができていて何ができていないのか要因を探しやすくなると考えている。喉頭の見本を見ることで、通常音声とは明らかに違う動きであることを理解できる。また、横隔膜の見本を見ることで、グロウル発声時の息を吐き出すスピードと量を自分と見比べることができる。

5. 制約と展望

体内を映し出す手段としてCTやMRIも存在するが、撮影に手間がかかるうえに、リアルタイムな提示が困難である。また、CTの場合は放射線を利用するため、身体的な負担が大きい。これらの観点から、本研究では身体的な負担が少なく手軽で、リアルタイムな提示ができる超音波映像を使用している。ただし、超音波映像は体内の状況によって歪みが生じる可能性がある。今後、超音波映像の歪みの補正機能の実装を検討していきたい。

本研究で使用した超音波映像装置(CMS600P2)は、1つの探触子からの入力しか処理できないため、喉頭と横隔膜の超音波映像を同時に提示できない。喉頭と横隔膜を同時に提示するためには、超音波映像装置を2台用意する必要がある。

超音波映像の見え方は使用する探触子の種類によって変わることがある。グロウルの見本と見比べるときは、同じ形状の探触子が使われているか確認し、違う形状の場合はその特性を考慮しながら観察する必要がある。

探触子を当てるときに男女間で喉仏の位置が違うことは考慮したい。男女差については男女両方の見本を用意することで対応する予定である。さらに細かい個人差については、認識が不可能になるほどの影響はないと考えている。なぜなら体内組織の相対的な位置関係は多くの人で共通しており、その知見が医療検査に利用されているからである。

超音波映像のAR表示は利用者によるグロウルの正誤判断だけでなく、発声の中間的な情報を学習できる点が特徴だと考えている。今後、超音波映像のAR表示が初心者のグロウルの学習において有用か検証を行ってきたい。

超音波映像のAR表示は、グロウル上級者の間で発展的な技術を共有する手段としても活用できる可能性がある。グロウルは感覚的な言葉で説明されることが多く、上級者

間の共通認識が形成されていない。マイク・カメラからのデータだけでなく、超音波映像も併せて共有することで、グロウルの技術をさらに細かい分類で体系化できると考えている。

6. まとめ

超音波映像を用いることでグロウルの学習を支援する手法を提案した。超音波映像装置は身体的な負担が少なく、リアルタイムに体内を観察できる。また、超音波映像と身体的位置の対応を分かりやすくするためにAR表示を用いている。声の生成については喉頭を、呼吸については横隔膜を観察することで、グロウルの学習を支援できる可能性がある。今後、本提案手法がグロウルの学習において有用か検証を行っていきたい。

参考文献

- [1] 森秀明, 平井都始子. はじめての超音波検査. 文光堂, pp.2-15, 2009.
- [2] 古川まどか, 古川政樹. 頭頸部エコーアトラス. 診断と治療社, pp.2-6, 2016.
- [3] 東義孝. パワーアップいまさら聞けない腹部エコーの基礎: DVDで学ぶ超音波検査. 学研メディカル秀潤社, p.302, 2011.
- [4] Oriol Nieto, Voice Transformations for Extreme Vocal Effects. POMPEU FABRA UNIVERSITY, 2008.
- [5] Philipp P. Caffier, Ahmed Ibrahim Nasr, Maria del Mar Ropero Rendon, Sascha Wienhausen, Eleanor Forbes, Wolfram Seidner, Tadeus Nawka, Common Vocal Effects and Partial Glottal Vibration in Professional Nonclassical Singers, Journal of Voice, Vol.32, No.3, pp.340-346, 2018.
- [6] Marco Guzman, Karol Acevedo, Fernando Leiva, Vasti Ortiz, Nicolas Hormazabal, Camilo Quezada. Aerodynamic Characteristics of Growl Voice and Reinforced Falsetto in Metal Singing, Journal of Voice, Vol.33, No.5, pp.803.e7-803.e13, 2019.
- [7] NOV. 地獄のボーカル・トレーニング・フレーズ. リットーミュージック, pp.110-111, 2010.
- [8] ビル・マーティン. ヘヴィロック唱法の奥義. リットーミュージック, pp.66-73, 2007.
- [9] 前掲. ヘヴィロック唱法の奥義. pp.170-171.
- [10] 後藤真孝, 齋藤毅, 中野倫靖, 藤原弘将. 歌声情報処理の最近の研究, 日本音響学会誌, Vol.64, No.10, pp.616-623, 2008.
- [11] Tomoyasu Nakano, Masataka Goto, Yuzuru Hiraga. MiruSinger: A Singing Skill Visualization Interface Using Real-Time Feedback and Music CD Recordings as Referential Data. Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007), Beijing, pp. 75-76, 2007.
- [12] 村尾忠廣. 視覚フィードバックによる音痴の治療—SINGADからVSG(trace)による方法へ—, 音声言語医学, Vol.41, No.3, pp.255-259, 2000.
- [13] Naoki Kimura, Michinari Kono, and Jun Rekimoto. SottoVoce: An Ultrasound Imaging-Based Silent Speech Interaction Using Deep Neural Networks. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 146, 1-11, 2019.