

波形の特徴に基づく鶏の鳴き声合成*

© 鳥居美月, 吉野夏樹, 矢田部浩平, 新村毅 (農工大)

1 はじめに

雛の成長には鳴き声による親子間コミュニケーションが重要だが、養鶏現場では親子分離飼育が行われるため、雛の恐怖心増大やそれに伴う成長阻害が問題視されている。恐怖心低減には、雛に向けた親鶏の鳴き声再生の有効性が示されている [1]。しかし、同じ音刺激を受け続けた雛は、その刺激に馴れて反応性が低下 (馴化) する。従って、雛の鳴き声に即座に応答するリアルタイム性と、馴化防止のために多様な鳴き声を合成できるシステムが求められる。そこで本稿では、採餌時などに親鶏が発する鳴き声 (food call) を対象とし、ソース・フィルタモデルに基づく軽量な合成法を提案する。本手法では乱数を用いることで多様な鳴き声の生成を実現する。また、主観評価実験により、合成した鳴き声が自然であることを確認した。

2 鶏の発声メカニズムと生成モデル

鳥類は、独自の発声器官である鳴管で音を生成し、気管やくちばしで調音された音を発している。鶏の鳴管は気管の分岐点に存在し、図-1 に示すような構造を持つ。発声時は、鳴管筋によって張力が調整された膜を呼気が振動させることで音が発生する。図-1 のように、鶏の鳴管には音の発生源が 3 か所存在するため、発生する音は非常に複雑である。また、発声時の呼気圧の動的な変動や筋肉による気流制御に起因して、この音には多様な振幅変調が生じる [3]。こうして発生した音は気管からくちばしに至る声道で調音される。鶏の気管は一樣な細長い音響管として近似でき、その長さが低周波帯域の共振の決定要因となる。また、開口端であるくちばしの開き具合は高次の共振特性に影響を及ぼす [4]。

この発声メカニズムは、音源で発生した音が共鳴器で調音されて出力されるとみなすソース・フィルタモデルとして捉えることができ、鳥類への適用が実証されている [2]。このモデルに基づき、本稿で扱う food call の共振特性を確認するため、雌の烏骨鶏から収録した food call に線形予測分析を適用しフィルタ特性を解析した。得られたフィルタの極のヒートマップを図-2 に示す。本図より、food call の共振周波数は共通の周波数帯域 (赤枠) に集中していることが分かる。これは、発声時に共鳴器の形状が概ね一定に保たれていることを示唆している。

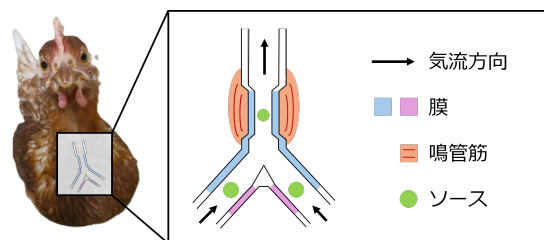


図-1 鶏の鳴管の位置 (左) と模式図 (右)

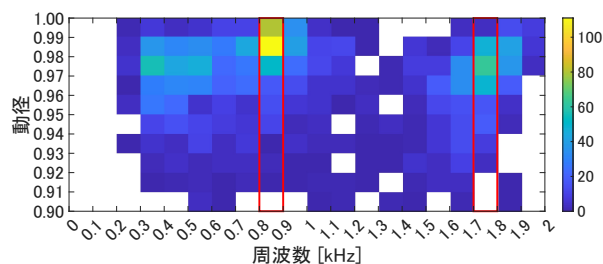


図-2 共振フィルタの極をまとめたヒートマップ。16 kHz でサンプリングされた 654 個のデータに対し、次数 8 の線形予測分析によって極を算出した。横軸は周波数 (幅 0.1 kHz)、縦軸は動径 (幅 0.01) を表す。

3 音源信号波形の特徴と合成法

鳴き声合成システムの実用化においては、馴化に対処するため多様な鳴き声を提示する必要がある。しかし、同一音声の反復を避けるために膨大な波形データを保持することは、メモリや計算コストの観点から望ましくない。従って、個別の録音データに依存しない合成法の構築が必要であり、そのためには音源信号の普遍的な特徴の特定が求められる。

3.1 音源信号推定

Food call の音源信号の特徴を確認するため、声道形状の不変性を仮定し、逆フィルタリングで音源信号を推定した。図-2 で極が集中した 0.85, 1.75 kHz に極を配置した共振フィルタを用いた。なお、約 0.3 kHz の極を含めると波形のモデル化が困難になるため除外した。図-3 に示す推定結果から、音源信号波形には時間的広がりを持つパルス状の成分が複数存在し、それらの振幅が動的に変化していることが確認できる。

これらの特徴のうち、合成に不可欠な成分を特定するため、事前検討を行った。その結果、振幅が大きいパルス成分が不可欠であるが、これを時間的広がりを持たない単純なパルスで代用すると、合成音の再現性が著しく低下することが判明した。従って、大きい振幅を持つパルスの時間的広がり、音の立ち下がった後

* Waveform-feature-based synthesis of chicken vocalization. By Mitsuki TORII, Natsuki YOSHINO, Kohei YATABE and Tsuyoshi SHIMMURA (Tokyo University of Agriculture and Technology).

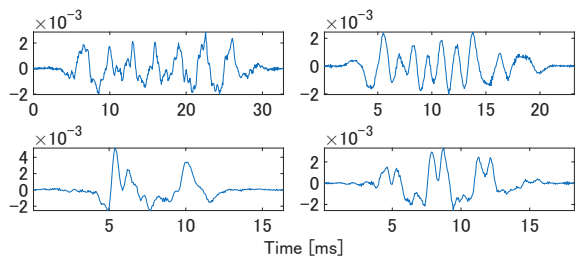


図-3 逆フィルタリングで推定した音源信号波形 4 種

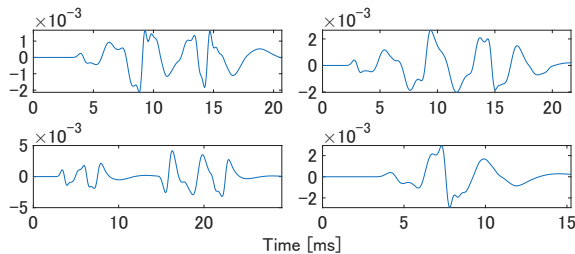


図-4 フィルタ適用後の生成した音源信号波形 4 種

から次の立ち上がりまでの時間間隔、及び振幅包絡を適切にモデル化することが合成において重要となる。

3.2 提案する合成法

この知見に基づき、本稿では多様な波形を軽量に生成するため、乱数を使用した音源信号生成法を提案する。パルス成分には有限長で打ち切ったガウシアンパルスを用い、乱数で決定した個数、幅、及び間隔でこれらを連結する。さらに、データから推定した音源信号の正負の包絡をランダムに選択し、隣接するパルス間で正負が交互に反転するように包絡を適用する。事前検討を踏まえ、パルス数は 2–10、幅は 0.625–2.5 ms、立ち下がりから立ち上がりまでの間隔は 0–0.125 ms の範囲の乱数で決定する。また、生成された波形に 6 kHz の正弦波による振幅変調を施すことで、気流の乱れや膜の不規則な振動などに起因する影響を反映させ、音響的な自然性の向上を図る。合成音に低周波及び高周波帯域の不要な振動成分が含まれないようにするため、カットオフ周波数 0.15 kHz のハイパスフィルタと 2 kHz のローパスフィルタを適用する。このようにして生成された信号に共振フィルタをかけて合成音を得る。

4 合成音評価

まず、生成された音源信号の多様性を確認するため、ローパス及びハイパスフィルタを適用した後の波形例を図-4 に示す。本図から、乱数の導入により同一の信号は生成されていないが、各パルスの形状が概ね一様であることが確認できる。このため、同じ振幅包絡を用いた場合に合成される鳴き声が類似してしまう懸念があり、今後は個々のパルス形状を変化可能な音源信号生成法の構築が望まれる。

表-1 主観評価実験におけるスコアと t 検定における p 値

データ群	スコア (平均値 $\pm$ 標準偏差)	p 値
録音データ	56.3 $\pm$ 29.6	2.5×10^{-5}
合成音	72.1 $\pm$ 23.3	

次に、合成音の自然さを定量的に評価するため、主観評価実験を行った。評価用音源として、録音データ 15 音源及び合成音 15 音源の計 30 音源を用意し、ヘッドホンで聴取した各音源の自然さを 1–100 点で評価した。評価の基準として録音データ 3 音源をあらかじめ提示し、これらを 100 点の自然な音と認識させた上で、0, 25, 50, 75, 100 点をそれぞれ「人工的・やや人工的・どちらでもない・やや自然・自然」として評価させた。なお、実験参加者には各評価音が録音データと合成音のどちらであることを提示しないブラインドテストとした。実験参加者 6 名が計 30 音源を評価した結果から、録音データと合成音それぞれについて得られた 90 試行分のスコアの平均値と標準偏差、及び両群間の t 検定における p 値を表-1 に示す。 t 検定の結果、両群には統計的に有意な差が認められ、録音データと比較して合成音の方が有意に高い評価を得た。この結果の要因として、環境雑音の有無が影響していると考えられる。録音データには収録環境に起因する無視できない程度の環境雑音が含まれる。一方、ソース・フィルタモデルに基づく本合成法は雑音を含まないクリアな音を合成できるため、主観的な自然さの評価が高くなったと考えられる。

5 むすび

本稿では、計算量が小さく多様な food call を合成可能な手法として、乱数を用いたソース・フィルタモデルに基づく合成法を提案した。合成音は録音データを上回る自然さを有することが確認できたが、音源信号の多様性に関しては改善の余地が残った。今後は、パルス形状の均一性を緩和して非周期的な波形を実現し、より多様な鳴き声の合成を図るとともに、合成音に対する雛の行動応答を観察し、恐怖心低減や馴化抑制効果を検証する。

参考文献

- [1] T. Shimmura, K. Hayakawa, K. Ichishima, Y. Tategaki, N. Nozaki, and R. O. Tachibana, “Chick-computer interaction using sounds,” *11th Int. Symp. Adapt. Motion Anim. Mach.*, pp. 134–135 (2023).
- [2] S. Nowicki, “Vocal tract resonances in oscine bird sound production: Evidence from birdsongs in a helium atmosphere,” *Nature*, **325**, 53–55 (1987).
- [3] C. P. H. Elemans, M. Muller, O. N. Larsen, and J. L. van Leeuwen, “Amplitude and frequency modulation control of sound production in a mechanical model of the avian syrinx,” *J. Exp. Biol.*, **212**, 1212–1224 (2009).
- [4] N. H. Fletcher and A. Tarnopolsky, “Acoustics of the avian vocal tract,” *J. Acoust. Soc. Am.*, **105**, 35–49 (1999).