電気式人工喉頭を用いた歌声の音質改善

姫野 剛至^a 屏 貴文^a 田中 裕人^a 水町 光徳^a 中藤 良久^a 松井 謙二^b

^a九州工業大学 ^b大阪工業大学

1. はじめに

現在、病気などで喉頭を摘出し発声することができなくなった喉頭摘出者が全国に約 3 万人おり、喉頭摘出者の多くは電気式人工喉頭(以下 EL)を利用している。喉頭摘出者は高齢者の割合が高く、歌を楽しみたいというニーズがあり、EL にも歌唱を行う機能(以下歌モード)が搭載されている。しかし、EL の歌モードはピッチを段階的に上げ下げするだけであり、人の歌声らしくないという問題がある。これまで EL 音声の音質改善の検討⁽¹⁻²⁾は行われてきたが、歌声における音質改善の取り組みはほとんど行われていないのが現状である。

本研究では、EL に使用される振動体の代わりに振動スピーカを使用し、歌声にはどのような声帯音源が適しているかを検討し、歌声の音質改善を図る。

2. 種々の擬似声帯音源による歌声比較

2.1 擬似声帯音源の作成

本稿では、擬似的に作成した声帯音源(擬似声帯音源)を用いた歌声の主観評価を行う。使用する歌声は、健常者 1 名が童謡「きらきら星」を以下の A~C の 3 種類の擬似声帯音源を用いて歌唱したものである。歌唱者はこれらの擬似声帯音源を振動スピーカから再生し、喉に押し当てて歌唱する。なお、D は健常者本人が自分の声で歌唱したものであり、比較の為に収録した。

- A. 通常の歌声から LPC 分析により抽出した擬似声帯音源(LPC 残差波)
- B. 音声合成等で用いられる Rosenberg 波
- C. EL の歌モード
- D. 生歌

A の LPC 残差波は、D の生歌から LPC 分析により抽出した包絡を用いて、逆フィルタリングすることによって生成した擬似声帯音源である。また、B の Rosenberg 波は主に音声合成において声帯振動として用いられており、式(1)を用いて作成する。

$$f(t) = \begin{cases} 3\left(\frac{t}{\tau_1}\right)^2 - 2\left(\frac{t}{\tau_1}\right)^3 & (0 \leq t \leq \tau_1) \\ 1 - \left(\frac{t-\tau_1}{\tau_2}\right)^2 & (\tau_1 \leq t \leq \tau_1 + \tau_2) \end{cases} \quad (1)$$

ここで $\tau_1 = 0.95$ $\tau_2 = 0.90$ とする。C の EL の歌モードはボタンを押すごとに、ピッチを段階的に上げ下げするものである。

2.2 聴取実験による歌声比較

A~D の 4 種の歌声について、シェッフェの対比較法を用いた聴取実験を行う。本実験は、きらきら星の中から「きらきらひかる」、「おそらのほしよ」、「まばたきしては」、「みんなをみてる」の 4 フレーズ用い、健常者 10 名を被験者とした。

実験結果を Fig. 1 と Fig. 2 に示す。Fig. 1 は、横軸が評価値であり、値が -2 に近づくほど自然性が高い(肉声に近い)ことを表している。また Fig. 2 は、2 種ずつの歌声を比較し、2 種の歌声に対してどの程度自然性に差があるかを表している。縦軸が嗜好度の差の値であり、負方向に大きくなると左の歌声が、正方向に大きくなると右の歌声が自然度が高いことを表す。なお、信頼区間は 95% である。Fig. 1 と Fig. 2 より、どのフレーズにおいても生歌が最も自然性が高く、続いて LPC 残差波、Rosenberg

Fig. 1. Preference score of each methods in 4 phrase.

Fig. 2. Comparison between the two methods (95% confidence interval.)

波、EL という順となった。このことから、LPC 残差波が最も自然性の高い擬似声帯音源であることが分かった。しかし、Fig. 1 より、擬似声帯音源による歌声と、生歌との間には以前大きな差があり、改善の余地があることが分かる。

3. LPC 残差波を用いた歌声の音響的分析

本章では、LPC 残差波が生歌と音響的にどのように違いがあるか分析を行った。

3.1 音圧の比較

生歌と LPC 残差波を用いた歌声の音圧での違いを分析した。Fig. 3 は、生歌と LPC 残差波の「きらきらひかる」の部分における音圧の波形である。Fig. 3 より、/ki/、/hi/ など、母音の /i/ の発声において平均約 20[dB] の差が見られた。また、/u/ や /o/ などの母音においても同様に約 10~20[dB] の差が見られた。これらの結果から、LPC 残差波では /i/、/u/、/o/ などの口を狭めて発声する母音において音圧が大きく下がると考えられる。

3.2 スペクトル包絡の比較

LPC 残差波を用いた歌声と生歌において、母音のスペクトル包絡にどのような違いが表れるか分析を行った。FFT の分析窓は 100 [ms] の 256 である。きらきら星の「きらきらひかる」「おそらのほしよ」「まばたきしては」「みんなをみてる」の 4 フレーズ中に含まれる母音を切り出し、平均を算出して分析を行った。Fig. 4 はその中の /i/ のスペクトル波形である。Fig. 4 より、生歌の /i/ は約 500~2000[Hz]、4000~5000[Hz] が谷になっているのに対し、LPC 残差波の /i/ は谷が見られず、生歌の /i/ のスペクトル包絡とは大きく異なる。また、/u/ の母音においても同様の結果が得られた。よって、LPC 残差波を用いた歌声において /i/ や /u/ の発音が正しく行えていないと考えられる。

3.3 基本周波数遷移の比較

基本周波数遷移にどのような違いが出るか比較を行った。Fig. 5 に LPC 残差波を用いた歌声と生歌の基本周波数遷移の波形を示す。Fig. 5 から、LPC 残差波を用いた歌声は、通常の歌声とほぼ同じ遷移をしていることが分かった。また、生歌の基本周波数を基準とすると、LPC 残差波と生歌は平均 3.67[Hz] 違いしかなく、LPC 残差波が生歌の基本周波数遷移とほぼ同じであることが分かった。

4. 終わりに

本稿では、EL に替わる擬似声帯音源による歌声について、聴取実験と音響的分析を行った。その結果、LPC 残差波が他の擬似声帯音源に対して、自然性が優れていた

Fig. 3. Sound pressure difference of the human voice and LPC of the phrase “kirakirahikaru”.

Fig. 4. Voice spectral comparison of vowel /i/ in the Twinkle Twinkle Little Star.

Fig. 5. Comparison of the fundamental frequency transition of the various voice.

ことが分かった。また、LPC 残差波は生歌に対して、/i/、/u/ の母音において大きく音圧が下がり、凹凸のないスペクトル包絡になることが分かった。

今後、より自然性の高い歌声の実現のために /i/ や /u/ を正しく発声する方法や、LPC 残差波の音圧やスペクトルの補正方法について検討する。

謝 辞

本研究の一部は、文部科学省学術研究助成基金 (NO.24500664) の援助を受けて行われた。

文 献

- (1) 粕谷英樹, 菊池義信:「韻律制御可能な電気式人工喉頭の現状と課題」, 音響講論, 1-2-3, 2014
- (2) 藪謙一郎, 伊福部達:「電気式人工喉頭のための拡声器付きウェアラブル装着具の試作」, 音響講論, 2-2-4, 2013