

## 視覚障害者のためのカラオケ支援システム Karaoke Support System for Visually Impaired Person

上倉 一人<sup>†</sup>      遠藤 勇佑<sup>†</sup>      石井 孝太郎<sup>†</sup>  
Kazuto Kamikura   Yusuke Endo      Kotaro Ishii

### 1. はじめに

音楽を気軽に楽しむことができるカラオケは、心身の健康増進に役立つことが知られている。しかしながら、視覚障害者はカラオケ映像に表示される歌詞が読めない、または読みづらいため、歌詞をすべて覚えているか、支援者に耳元で先読みしてもらう必要がある。これに対して歌詞をカラオケと連動させて点字ディスプレイに表示させるシステム[1]が開発されているが、点字が読めないと利用できない。筆者らは上記の課題を解決するために、視覚障害者が晴眼者と同じように一人でカラオケを楽しむことができる支援システムの実現を目指している。本稿では本カラオケ支援システムの全体イメージについて述べ、その中の技術課題として、歌詞先読みの開始タイミングと先読みスピード、カラオケ曲タイトルの特定、歌詞先読み音声とカラオケとの同期再生、の各々に関する検討状況について述べる。

### 2. カラオケ支援システムの全体イメージ

カラオケ支援システムの全体イメージを図 1 に示す。まず、①複数のカラオケ曲に対して歌詞先読み音声ファイルを音声合成により作成し、それらのファイルを小型コンピュータに保存しておく。カラオケ曲を歌う際には、②カラオケ画面に表示された曲タイトルを文字認識によって特定し、③特定した曲に対応する歌詞先読み音声ファイルをイントロ開始と同時に再生する。その音声をオープンイヤ型ワイヤレスイヤホンに伝えることによって、適切なタイミングで歌詞を利用者に伝える。

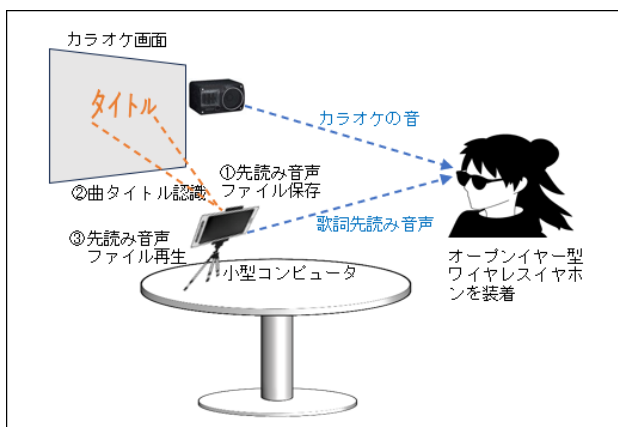


図 1 カラオケ支援システムの全体イメージ

本システムにおける技術課題として以下の 3 つが挙げられる。1 つ目は歌詞先読み音声ファイルの作成方法である。どのようなタイミングで先読みを開始し、どのようなスピードで先読みすれば最も歌いやすいのか検討する必要がある。2 つ目は曲タイトルの認識である。カラオケ画面の中から曲タイトルを高精度で認識することが必要である。3 つ目は歌詞先読み音声とカラオケとの同期再生である。歌詞先読み音声がかラオケに合わせて再生されることが必要である。

### 3. 歌詞先読みの開始タイミングと先読みスピード

カラオケ映像では、歌う前に 2 行程度の歌詞が同時に表示され、歌うタイミングに合わせて一文字ごとに文字色に変化する。既に実施した検討[2]では、歌詞先読みの開始タイミングはカラオケ画面に歌詞が表示されると同時とするのが適切という結果が得られた。また先読みスピードについては、聞き取りやすさを考慮し通常よりやや速い一定速度とした。上記の条件で 4 曲について評価実験を行った結果、BPM(Beats Per Minute)が 80~110 程度の曲については「歌いやすい」、「歌うことができる」という評価だったが BPM が 160 を超える速いテンポの曲については「歌いにくい」という評価となった。理由として、歌うタイミングと歌詞先読み音声が重なることが多く、歌詞が聞き取りづらいたことが挙げられた。そこで今回は、読み上げ速度を[2]で設定した速度に対して 1 倍、1.5 倍、2 倍の 3 種類として評価実験を行った。被験者は 20 代の晴眼者 4 名であり、歌詞を見れば歌うことができる程度に曲を覚えている。評価尺度は以下のとおりである。

- 5: 完全に歌えた
- 4: (5 と 3 の中間)
- 3: 8 割以上歌えた
- 2: 5 割以上 8 割未満歌えた
- 1: 5 割未満しか歌えなかった

結果を表 1 に示す。数値は 4 名の評点の平均である。表 1 より、3 曲とも 1 倍速が最も評点が高く、速度が上がるほど評点が低くなっていることが分かる。特に天体観測 (BPM: 約 168) については[2]での評価結果と異なり、1 倍速でもほぼ歌えるという結果になった。理由として、この曲に対する被験者の経験度の違いが考えられる。この点については今後の評価方法を見直す必要があるといえる。また、どちらも晴眼者を被験者としているが、視覚障害者と晴眼者では速聴能力が異なる可能性があるため、今後は視覚障害者を対象とした評価実験も行う予定である。

<sup>†</sup> 東京工芸大学 Tokyo Polytechnic University

表 1 評価実験結果

曲タイトル	1 倍速	1.5 倍速	2 倍速
チェリー (BPM : 約 97)	4. 5	3. 8	2. 5
世界に一つだけの花 (BPM : 約 99)	4. 5	3. 8	2. 5
天体観測 (BPM : 約 168)	4. 5	3. 5	2. 3

#### 4. 曲タイトルの特定

小型コンピュータに付属または外付けされたカメラでカラオケ画面を撮影し、OCR(Optical Character Recognition)により文字認識することで曲タイトルを特定する。現時点では、実験において 14 曲のタイトルに対し 11 曲を 100%の精度で認識できた PaddleOCR を用いている。100%の精度で認識できなかった 3 曲に対して、前処理として明度とコントラストを変化させたが、認識精度を上げるための共通の処理を定めるには至っていない。

今回実験に用いた 14 曲の画像はモニタの中心とほぼ同じ高さで撮影したが、実運用を考慮すると利用者が細かい点まで意識してカメラ設置することは困難である。したがって今後は、カメラの設置位置や照明の違いなどを考慮した検討も行っていく必要がある。

#### 5. 歌詞先読み音声とカラオケとの同期再生

歌詞先読み音声ファイルのカラオケに同期させて再生する。その方法として、カラオケのイントロに相当する時間だけ音声ファイルの冒頭に空白時間を設けておき、イントロが開始すると同時に音声ファイルの再生を開始することを考えている。そのためには、小型コンピュータのマイクに入力する音の大きさや周波数成分などから、イントロの開始部分を検出することが必要となる。この点について今後検討を進めていく。

#### 6. むすび

視覚障害者のためのカラオケ支援システムについて、その全体イメージを紹介した。また、本システムを実現するに当たって解決すべき技術課題の検討状況と今後の予定について述べた。

なお、現状は歌詞先読み音声ファイルを手動で作成しているが、2 行程度の歌詞表示タイミングや文字色変化のタイミングなどを利用して自動で作成することも、今後検討していきたい。

#### 謝辞

本研究を進めるに当たって助言をいただいた JSU 支援技術研究所 代表 (元 公益財団法人 新産業創造研究機構 技術アドバイザー) 内原正一氏に感謝する。

#### 参考文献

- [1] “点字カラオケ”, [https://www.nippontelesoft.com/menu/disp\\_005.html](https://www.nippontelesoft.com/menu/disp_005.html)
- [2] 遠藤, 上倉, “カラオケ支援における歌詞先読みタイミングに関する一検討”, 2024 年映像情報メディア学会年次大会, 21B-4 (2024).