

令和 5 年 6 月 26 日現在

機関番号：11401

研究種目：基盤研究(C) (一般)

研究期間：2017～2022

課題番号：17K01553

研究課題名(和文) 視覚障害者の映像鑑賞における音声合成利用の可能性と新展開

研究課題名(英文) Possibilities and Developments in the Use of Speech Synthesis for People with Visual Impairments to Enjoy Movies

研究代表者

中島 佐和子(Nakajima, Sawako)

秋田大学・理工学研究科・講師

研究者番号：40453542

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：超高齢社会にあって、視覚や聴覚に障害を有する方々への映画や映像のバリアフリー化が急務である。この世界的動向に対し本研究では、音声合成や話速変換技術を活かすことにより、視覚障害者の映像鑑賞体験を充実させ、音声ガイド普及を促す道筋を探った。視覚障害者を対象とした肉声と合成の音声ガイドの比較実験から、合成音声には、話速を高速にしても聞き取れる、また、話速を高速にしても違和感が少ないなどの優れた点が得られた。これらの利点を活かし、合成音声を用いた音声ガイド制作を加速させる必要がある。そこで、音声合成を用いた音声ガイド制作ソフトを開発し、実証実験を通じて課題と可能性を得た。

研究成果の学術的意義や社会的意義

視覚障害者を対象とした映画の音声ガイドの聞き取り実験、および、調査からは、視覚障害者の聴覚特性に関する基礎的知見を得ることができた。視覚障害者が日常生活を送る上で合成音声による情報収集は欠かせない。視覚障害者の感覚や意見からは、今後、社会が合成音声などの技術をどのように活用していくか、新しい音声表現の可能性を得る上で重要な視点を得られた。また、音声合成を用いた音声ガイド制作ソフトは、実証実験を通じて課題を抽出し、また、改善したことで、音声ガイド制作現場で利用価値の高い技術となった。

研究成果の概要(英文)：In response to this global trend, this study explored ways to improve the movie experience for people with visual impairments and to promote the use of audio description (AD) through the use of speech synthesis and speech rate conversion technologies. Comparative experiments between human AD and synthesized AD showed that synthesized AD has advantages such as being audible and causing little discomfort even at high speech rates. It is necessary to exploit these advantages of synthesized speech and accelerate the production of AD. Therefore, AD production software using speech synthesis has been developed, and problems and opportunities have been identified through demonstration experiments.

研究分野：福祉情報工学

キーワード：視覚障害者支援 映画・映像 音声ガイド 音声合成 ヒューマンインタフェース

1. 研究開始当初の背景

日常生活を豊かにし楽しく文化的な活動を行う上で映画やテレビ鑑賞は欠かせない。一方で、加速する高齢者人口の増加を背景に様々な身体特性を有する人々（視覚障害や聴覚障害等）の割合が増加している。映画鑑賞のバリアフリー化への障害当事者からの期待は非常に高く、高齢者や障害のある人たちへの情報保障という観点も含めて、その実現がいち早く望まれる。しかし、普及面では課題が多い。映画の音声ガイドに至っては、平成24年の邦画とアニメの劇場公開作品の音声ガイド付与率は1.1%、DVDおよびBlue-rayでは0.1%に過ぎない。様々な角度から音声ガイド普及に繋がる取り組みが必要な状況にある。

そのような状況の中、音声ガイドや字幕を映画館で本編に合わせて同期提示するシステムの実用化など、環境整備も進められている。とりわけ、この同期提示システムを利用し、映画館でスマートフォンやタブレット端末を経由して、イヤホンやヘッドホンにより音声ガイドを聞き取るシステムの当事者評価は非常に高く、これまでの映画鑑賞スタイルを変えつつある。同システムを用いて申請者らが実施した試写会調査では、『セリフを聞くのと音声ガイドを聞くのとは聞き方が変わるため、イヤホンならどちらにも集中できる』など、利便性だけでなく聞こえに対する効果も確認できた。さらに調査を進めることで、セリフ音声を含む映画や映像の本編音響と音声ガイドの質的違いを聞き分け、それらを手がかりにそれぞれの音が持つ役割を識別し、情報の中身を上手く重ね合わせていることがわかった。

2. 研究の目的

そこで本研究では、音声ガイドに合成音声を用いることで、映画の登場人物のセリフや環境音や音楽と音声ガイドが持つ声質や音質の違いをより識別し易くすることで、音により表現できる世界を広げることができないかと考えた。合成音声の利点を広げることができれば、音声ガイド制作コストの削減により音声ガイドの普及にも繋がる。音声合成や話速変換技術を活かすことにより、視覚障害者の映像鑑賞体験を充実させ、音声ガイド普及を促す道筋を探ることを目指した。さらに、音声ガイド制作支援に対しても音声合成技術の適用を図り、総じて、視覚障害者の映像鑑賞環境を改善することで、豊かな文化的土壌を培うことを目的とした。

ここでは主に以下の2点に関して実施した研究成果についてまとめる。

- (1) 音声合成を用いた音声ガイド制作の課題と音声ガイド制作ソフトウェアの開発
- (2) 肉声および合成音声の映画音声ガイドの最適話速の評価

3. 研究の方法

(1)

112分の日本語の劇映画に対して、プロの音声ガイド制作者1名により音声ガイド台本を制作した。それをもとに、人間のナレーターによる音声ガイドの収録と音声合成による音声ガイドの2つを制作した。音声ガイドが公開されるまでの制作過程を記録し、合成音声を用いた音声ガイド制作の課題を抽出した。さらに、得られた課題や音声ガイド制作現場からのニーズを調査した上で、音声合成による音声ガイド制作ソフトを開発した。また、開発したソフトを用いて、5名のプロの音声ガイド制作者によって、映画予告編（1分26秒）の音声ガイドを制作した。さらに、予告編の本編映画のプロデューサーから、総合、挿入タイミング、話速、音量の4要素について制作された5つの音声ガイドの評価を得た。

(2)

視覚障害者8名、晴眼者8名を対象に、5秒程度の短い劇映画のワンシーンを用いて、シーンに付与された音声ガイドを7.42から18.56 moras/secに変化させた合成音声および肉声音声の音声ガイドを作成した。それらを用いて、「音声ガイドのみ提示する場合」と「映画音と音声ガイドを提示する場合」について階段法により最適な話速を得た。

4. 研究成果

(1)

はじめに、音声ガイドに合成音声を用いた場合の課題を把握するために、汎用の映画再生ソフトとExcelの表計算ソフトを用いて、約2時間の日本の劇映画を対象に、合成音声による音声ガイドとナレーターによる肉声の音声ガイドを制作する実証実験を実施した。その結果、合成の音声ガイドは、挿入タイミング、イントネーション、音量、間（ポーズ）について、映画プロデューサー、音響エンジニア、音声ガイド制作者によりスタジオにて約5時間20分の修正作業を必要とした。また、修正の後に完成した合成の音声ガイドの質は、ナレーターによる音声ガイドの約75%にしか達しなかった。

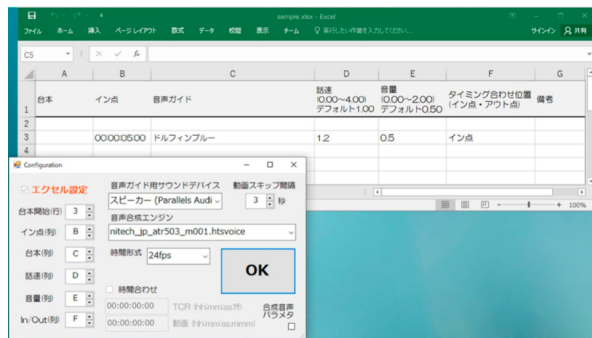
そこで、本研究ではC#を用いて図1に示す音声ガイド制作ソフトを開発した。メインウィンドウの上部には動画が、下部(動画領域の下)には、音声ガイドを付与させたい動画と音声ガイド

の音声波形が表示される。音声ガイドの波形領域を右クリックするとコンテキストメニューが表示され、「新規作成」をクリックすると音声ガイドの新規作成ウィンドウが表示され、音声ガイドの内容、話速、音量を入力することで音声合成され、右クリックした位置に音声ガイドが新規挿入される。挿入された音声ガイドは図 1(b)(c)のように水色の四角でハイライトされ、その四角をドラッグアンドドロップで時間軸に並行移動することで、音声ガイドの挿入位置を調整できる。また、作成された音声ガイドを編集、または、削除したい場合は、該当する音声ガイドの水色四角領域を右クリックすることで「編集」、「削除」メニューが加わったコンテキストメニューが表示され、そこから操作が可能である。図 1(b)右に音声ガイド台本の Excel ウィンドウを示す。音声ガイド制作ソフトのメインウィンドウ (図 1(b)左) で挿入した音声ガイドの音声ガイドの内容、話速、音量、挿入タイミングは即座に Excel に反映される。映画音と音声ガイド全体の音量バランスは、それぞれを 1 から 100 で設定することができる。

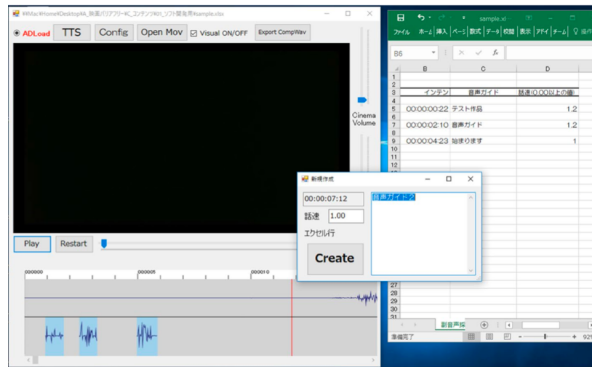
本ソフトを用いて、プロの音声ガイド制作者による映画予告編の音声ガイド制作実験を実施した。実験の結果、映画プロデューサーや音響エンジニアによる音声ガイド修正を経なくても、先に制作した合成の音声ガイドよりも 5%ほど質を向上させることができた。しかし、これまではナレーターの役割であった音声ガイドの話速や音量などのパラメータ設定を音声ガイド制作者が負担することには変わりなく、それらの要素の質には大きな課題が残った。

(2)

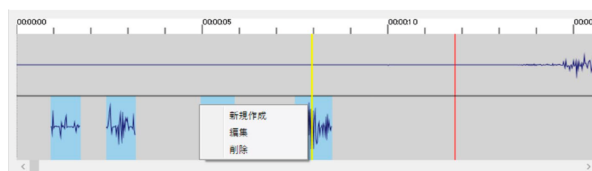
そこで本研究では、音声ガイド制作者の制作指針を得るために、視覚障害者と晴眼者を対象とした、映画音声ガイドの最適話速を評価した。実験の結果、「音声ガイドのみ提示する場合」は、視覚障害者では、肉声音声より合成音声の最適話速が高くなる傾向であったが、合成音声の最適話速はばらつきが大きく有意な差ではなかった。対して、晴眼者は合成音声の最適話速が有意に高い結果となった。一方、「映画音と音声ガイドを提示する場合」については、視覚障害者の最適話速の幅は狭まり、視覚障害者、晴眼者ともに有意に合成音声の方が最適話速が高い結果となった (図 2)。また、肉声音声と合成音声の最適話速の間に有意に強い正の相関が示された (図 3)。以上の結果から、映画音声ガイドの聞き取りやすい最適な話速は、映画本編の音に影響を受け一定の範囲に狭まる。しかし、合成音声の音声ガイドの最適話速の方が肉声の音声ガイドより早いバランスを保持したまま、個人ごとに異なる声質によらない最適話速に収まる。さらに、合成の音声ガイドについて、「映画音と音声ガイドを提示した



(a) Example of selected Excel spreadsheet opened via the software and configuration window



(b) Example of video-display and audio-waveforms areas and sub-window for creating new AD



(c) Example of film- and AD-waveforms areas and menu for creating, editing, and deleting AD opened beside yellow vertical line shown on right-click point

図 1 音声合成を用いた音声ガイド制作ソフト

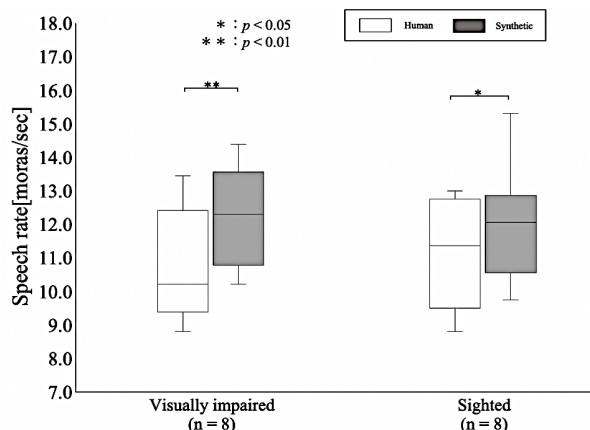


図 2 映画音を提示した場合の音声ガイドの最適話速

試行」における、視覚障害者が選択した音声ガイドと選択しなかった音声ガイドの音響的違いを分析したところ、音声ガイドを映画の環境音に重ねない工夫が見られた。図 4 上に視覚障害者群で選択されなかった話速 9.28 moras/sec, 同図下に視覚障害者群で選択された最適話速 10.21 moras/sec を示す。縦軸と横軸はそれぞれ音声の周波数と提示時間である。音声ガイドの終了時間に着目すると同図上の話速 9.28 moras/sec は 4.0 秒を超えて終了しているのに対して、同図下の話速 10.21 moras/sec は 4.0 秒より前で終了していることが分かる。本実験で提示した映画音は登場人物の一人が主人公の腕を 4 回叩く音が収録されている。4.0 秒付近には 4 つ目の叩く音があり、同図上の 9.28 moras/sec の音声ガイドは 4 つ目の叩く音と重なるのに対し同図下の 10.21 moras/sec の音声ガイドは重ならないことがわかった。以上のことから、合成音声を用いた音声ガイド制作で新たな課題となった話速設計を、肉声にはない合成音声の速聴の可能性という観点から捉え直すことができた。なお、視覚障害者への調査においても、「肉声に比べて合成音声は話速を早くしても違和感を生じにくい」という意見も得ている。

本研究では、さらにこれらの結果に基づき、(1) で開発した音声合成による音声ガイド制作ソフトの改良、ソフトを用いた初心者の音声ガイド制作実験と制作課題の検出方法の検討、また、視覚障害者による音声ガイドの話速調整機能の試作と評価などを実施し、合成音声の良さを引き出す話速設定方法や音量設定方法について検討した。以上より、視覚障害者の映像鑑賞体験を充実させるための新たな視点の提案とシステム開発の一部を实践できた。

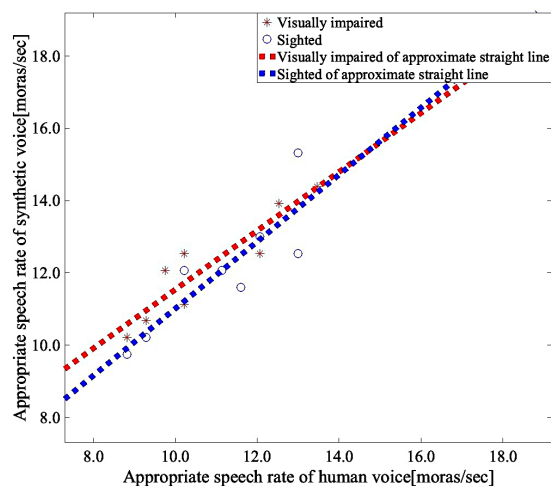


図 3 映画音を提示した場合の肉声と合成の音声ガイドの最適話速の相関

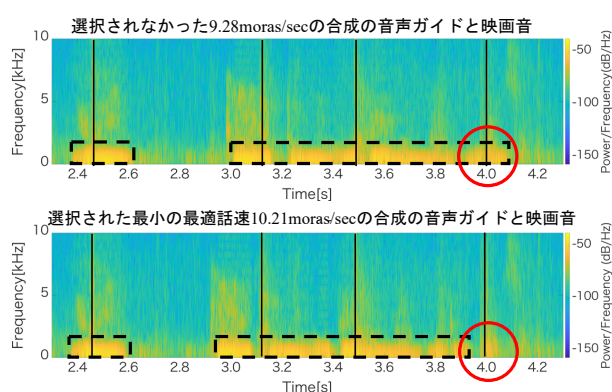


図 4 合成音声を用いた音声ガイドと映画音の関係

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Nakajima Sawako, Mitobe Kazutaka	4. 巻 --
2. 論文標題 Novel software for producing audio description based on speech synthesis enables cost reduction without sacrificing quality	5. 発行年 2022年
3. 雑誌名 Universal Access in the Information Society	6. 最初と最後の頁 --
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10209-022-00873-z	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Sawako Nakajima, Naoyuki Okochi, Kazutaka Mitobe and Tetsujiro Yamagami	4. 巻 -
2. 論文標題 Preliminary Study on Auditory Perception of Multiple Speech Synthesis with Different Speech Rates to Help Visually Impaired Persons Enjoy Movies	5. 発行年 2017年
3. 雑誌名 The Proceedings of the 8th International Conference on Materials Engineering for Resources (ICMR2017)	6. 最初と最後の頁 415-418
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 佐々木 怜衣, 中島 佐和子, 高橋 壱成, 大河内 直之, 山上 徹二郎, 水戸部 一孝	4. 巻 117(502)
2. 論文標題 視覚障害者の映画鑑賞支援を目的とした合成音声による音声ガイドの提示タイミングと心理的作用の評価	5. 発行年 2017年
3. 雑誌名 電子情報通信学会技術研究報告 (IEICE technical report : 信学技報)	6. 最初と最後の頁 103-107
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 直野剛也, 中島佐和子, 水戸部一孝	4. 巻 121(52)
2. 論文標題 映画音声ガイドの聞き取りやすい合成音声の話速と環境音の関係	5. 発行年 2021年
3. 雑誌名 電子情報通信学会技術研究報告 (IEICE technical report : 信学技報)	6. 最初と最後の頁 38-42
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 栗山稜平, 中島佐和子, 水戸部一孝	4. 巻 121(52)
2. 論文標題 音声波形とサムネイル表示を用いた音声ガイド制作アプリケーションの試作	5. 発行年 2021年
3. 雑誌名 電子情報通信学会技術研究報告 (IEICE technical report : 信学技報)	6. 最初と最後の頁 43-47
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 栗山稜平, 中島佐和子, 水戸部一孝	4. 巻 122(295)
2. 論文標題 映画の音声ガイド制作における経験者と初心者の差異	5. 発行年 2022年
3. 雑誌名 電子情報通信学会技術研究報告 (IEICE technical report : 信学技報)	6. 最初と最後の頁 41-46
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

[学会発表] 計9件 (うち招待講演 1件 / うち国際学会 1件)

1. 発表者名 直野剛也, 中島佐和子, 水戸部一孝
2. 発表標題 映画音声ガイドの聞き取りやすい合成音声の話速と環境音の関係
3. 学会等名 第112回 福祉情報工学研究会
4. 発表年 2021年

1. 発表者名 栗山稜平, 中島佐和子, 水戸部一孝
2. 発表標題 音声波形とサムネイル表示を用いた音声ガイド制作アプリケーションの試作
3. 学会等名 第112回 福祉情報工学研究会
4. 発表年 2021年

1. 発表者名 森田 宏幸, 佐々木 亜希子, 大河内直之
2. 発表標題 スペシャル鼎談「みんなで楽しむバリアフリー映画の魅力語る」－「猫の恩返しバリアフリー版製作の現場をご紹介します」－
3. 学会等名 DPI日本会議（招待講演）
4. 発表年 2020年

1. 発表者名 佐々木怜衣, 中島佐和子, 高橋孝成, 大河内直之, 山上徹二郎, 水戸部一孝
2. 発表標題 合成音声を用いた映像鑑賞のための音声ガイドの提示方法が与える心理作用について
3. 学会等名 平成30年度情報処理学会東北支部研究会
4. 発表年 2018年

1. 発表者名 Sawako Nakajima, Naoyuki Okochi, Kazutaka Mitobe and Tetsujiro Yamagami
2. 発表標題 Preliminary Study on Auditory Perception of Multiple Speech Synthesis with Different Speech Rates to Help Visually Impaired Persons Enjoy Movies
3. 学会等名 The 8th International Conference on Materials Engineering for Resources (ICMR2017) (国際学会)
4. 発表年 2017年

1. 発表者名 佐々木 怜衣, 中島 佐和子, 高橋 孝成, 大河内 直之, 山上 徹二郎, 水戸部 一孝
2. 発表標題 視覚障害者の映画鑑賞支援を目的とした合成音声による音声ガイドの提示タイミングと心理的作用の評価
3. 学会等名 第94回福祉情報工学研究会(WIT)
4. 発表年 2017年

1. 発表者名 直野剛也, 中島佐和子, 水戸部一孝
2. 発表標題 聞き取りおよび印象に及ぼす映画音声ガイドの話速の検討
3. 学会等名 情報処理学会東北支部研究会
4. 発表年 2022年

1. 発表者名 栗山稜平, 中島佐和子, 水戸部一孝
2. 発表標題 映画との音量バランスに着目した音声ガイド制作支援手法の提案
3. 学会等名 情報処理学会東北支部研究会
4. 発表年 2022年

1. 発表者名 栗山稜平, 中島佐和子, 水戸部一孝
2. 発表標題 映画の音声ガイド制作における経験者と初心者の差異
3. 学会等名 第118回福祉情報工学研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	大河内 直之 (Okochi Naoyuki) (30361679)	東京大学・先端科学技術研究センター・特任研究員 (12601)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	水戸部 一孝 (Mitobe Kazutaka) (60282159)	秋田大学・理工学研究科・教授 (11401)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関