

# 視覚障がい者向け音声情報スキミングのための 高速再生技術の高度化及び受聴支援装置に関する研究開発

視覚障がい者が超高速に音声をサーチしながら情報把握可能な録音図書や電子書籍再生機の開発  
【平成 23 年度助成事業】

## 研究開発事業の概要と背景

視覚障がい者は小説や雑誌などの書籍は朗読ボランティアにより製作された録音図書コンテンツを聞くことで情報を得ている。録音図書は通常の再生速度で聞くと多くの時間を要してしまうため、多くの視覚障がい者は2～3倍速で聞いているのが実情である。しかし3倍速近傍では意味理解が困難となり、また長時間聞きつづけると疲労が増してしまうため、3倍速以上の高速再生時にも内容把握ができ、楽に聞くことが可能な音声高速再生技術開発への強い要求がある。

平成 20 年度から三カ年の NICT 助成事業により、従来技術では 3 倍速程度が限界といわれていたものが、提案法によれば 3 倍速でも非常に明瞭に疲労感少なく受聴することを可能とし、また 5 倍速程度でも内容把握が可能となった。しかし視覚障がい者からは、おおよその内容が把握出来れば良い場合もあり、さらなる高速音声再生への要求が寄せられており、平成 23 年度からの二カ年計画において、音声情報スキミングのための高速再生技術の研究開発に取り組むこととした。

## 超高速音声再生技術の開発

本研究開発における基本技術は、NHK放送技術研究所から技術移転を受けた「話速変換技術」であり、以下の課題について研究開発を行った。

### 1) 概要把握を効率よく行うための5倍速以上の超高速音声再生技術開発

これまでは音源情報の欠落なしに5倍速までの高速音声再生アルゴリズムを確立したが、本研究開発では情報欠落を積極的に行うことにより5倍速以上の超高速音声再生によるスキミング受聴を実現している。音源情報の一部欠落を伴う場合、音源のどの部分を残し、どの部分を欠落(削除)させるかが重要となる。入力音源を解析して得られる韻律パラメータ(ピッチ、パワー、ポーズ長等の動的变化)から、再生速度に応じて欠落させる部分の判定を行い、欠落によって生じた「時間的な余裕」を活用することで、5倍速以上の超高速再生時においても音声コンテンツの概要把握を可能としている。また、言語別では日本語ではモーラ等時性を活かした高速再生技術、英語圏ではストレス間等時性を活かした高速再生技術についても実用化の検討を行った。韻律パラメータと言語的な重要性の関係抽出については、NHK放送技術研究所の協力を得て行った。

### 2) 国内外での普及を目指した多言語適用への最適パラメータ抽出および自動言語性適応

日・英・独語については適応パラメータの最適値を取得済みであり、さらに仏語、西語、中国語における高速音声再生時の最適パラメータの値を評価実験により求めた。また、音源情報から言語性識別を自動的に行い、高速音声再生対象言語への適応アルゴリズムの基礎検討を行った。

### 3) テキスト読み上げTTS(Text-To-Speech)への適応

テキストDAISYではTTSエンジンによる読み上げが行われるが、人間の発話とは異なる不自然な韻律となることが多いため、テキスト解析技術を付加するなどTTSエンジンによる音声に特化した新たな高速音声再生アルゴリズムの基礎検討を行った。韻律に関する指導は東京外国語大学およびNHK放送技術研究所の協力を得て行った。

## 事業化の状況

市販の携帯型 DAISY プレーヤー機器への話速変換ライブラリの実装について検討を行ったが、GPU性能、実行処理速度、メモリ空間などの制約が大きく実現に至らなかった。そこで、話速変換ライブラリを組み込んだウィンドウズ版 DAISY プレーヤーを試作し、日本、アメリカ在住の視覚障がい者の方々に操作性などについて評価いただいた。

タブレット端末やスマートフォンのアプリケーションとして DAISY プレーヤーを開発している企業より話速変換ライブラリの組み込みについて引き合いがあり、具体化に向けて検討を進めている。

## 今後の展開

TTSエンジンからの音声出力への高速音声再生技術の適応は、一般的な電子書籍プレーヤーにも適応可能である。電子書籍規格とDAISY規格は類似しており、視覚障がい者のみならず一般人にも有効な技術となり、家庭における録画再生機、ICレコーダ、語学学習装置などへの応用が可能である。

## 事業実施データ

### ○財団法人NHKエンジニアリングサービス (東京都)

(共同研究機関: アダプティブテクノロジー)

○学会発表 2 件