

論 文 要 旨

Abstract

論 文 題 目

Title Myanmar Text-To-Speech System with Rule-based Tone Synthesis
(ミャンマー語の声調の規則合成とテキスト音声変換システム)

The MyanmarTTS system is a rule-based speech synthesis system, in which fundamental speech units are demi-syllables with Level tone. We use Myanmar (Burmese) official language for the TTS system. To construct the TTS system, monosyllabic words are analyzed and the parameters are obtained for synthesis of Myanmar tones. This TTS system uses a source filter model and furthermore a Log Magnitude Approximation Filter is employed for the vocal tract character. Myanmar is a tonal language comprising 4 different tones. The Level and High falling tones belong to the long-tone group and the Falling and Checked tones belong to the short-tone group. Checked tone is accompanied by a glottal stop. .

In this study, normalization method is different from other TTS systems such as Thai and Vietnamese. The synthesized tones are evaluated by employing listening tests. These results show that our rule-based linear pattern gives high intelligibility for the synthetic sound and is considered to be sufficient for MyanmarTTS system. The high intelligibility of synthesized tone was confirmed through listening tests of synthesized words with correct rates 95.6% for male and 97.8% for female speech. As a result, we showed that linear pattern is sufficient for Myanmar tone synthesis. The introduced normalization method is applicable for other tone synthesis rule of other tonal languages.

Large variability exists in the F_0 and the length uttered by different speakers and different syllables. Hence for speech synthesis normalization of F_0 and lengths are important to discriminate tones. Here, we implement tone rule by using two parameters; optimized F_0 and length. As an advantage in the proposed method, the optimized parameters can be separated to male and female group. The effectiveness of the proposed method is confirmed by the distribution of F_0 and length. Listening tests with high correct rates approve intelligibility of synthetic sound.

Furthermore, in the proposed method, the optimized parameters can be separated into male and female groups. The introduced proposed method is applicable for other tone synthesis rule of other tonal languages.

Name Win, Kyawt Yin

(様式5-2) 課程博士

平成23年2月17日

琉球大学大学院

理工学研究科長 殿

論文審査員

主査 高良富夫

副査 宮城隼夫

副査 M. R. Asharif



学位（博士）論文及び最終試験の終了報告書

学位（博士）の申請に対し、学位論文の審査及び最終試験を終了したので、下記の通り報告します。

記

申請者	専攻名 総合知能工学専攻 氏名 Win Kyawt Yin 学籍番号 [REDACTED]
指導教員名	高良富夫
成績評価	学位論文 ○合格 不合格 最終試験 ○合格 不合格
論文題目	Myanmar Text-To-Speech System with Rule-based Tone Synthesis ミャンマー語の声調の規則合成とテキスト音声変換システム
審査要旨（2000字以内）	

審査要旨：

本論文はミャンマー語のテキスト音声合成 (TTS) システムに関するものである。本システムでは Level 声調の半音節を音声合成単位としている。システムを構成するため、音節単語を分析している。この TTS システムでは、ソース・フィルタ・モデルを採用している。声道特性を表わすフィルタとしては、対数振幅特性近似フィルタが採用されている。ミャンマー語は 4 つの声調を持つ声調言語である。Level 声調と High Falling 声調は長い声調であり、Falling 声調と Checked 声調は短い声調に属する。Checked 声調は声門閉鎖を伴っている。

本システムを評価するため、ミャンマー語の音節の正聴率評価実験が行われている。その結果、正聴率は、92.6%となった。これは同じシステムを用いて日本語の音節を評価したものに比較して勝るとも劣らないものである。日本語より多様な音素を持つミャンマー語でも、対数振幅特性近似フィルタが十分な音節正聴率を示すことが示された。

本研究では、他の声調言語であるタイ語やベトナム語とは異なる正規化法が使用されている。正規化法は、特徴パラメータとして音節の長ささと音節中央部の基本周波数 (F0) 値を用いるものである。まず、これらのパラメータの分布が調べられた。その結果、正規化を行うとパラメータの分散が小さくなることが示された。

次に、合成された声調音声を、聴取実験で評価している。聴取実験結果によれば、本規則合成システムによる声調の F0 パターンは十分な正聴率を有しており、ミャンマー語の TTS システムとしては十分な性能であると考えられる。高い正聴率は聴取実験で確かめられている。男声の場合は 95.6%であり、女声の場合は 97.8%であった。この実験で、F0 の直線パターンが十分な性能を持っていることが分かった。提案されている正規化法は、他の声調言語の声調合成へも適用可能である。

異なる話者と異なる音節で、F0 パターンは大きく異なる。本研究では、F0 と長さに関して新たに性ごとに最適化を行うことを提案している。提案された最適化法の有効性は、F0 と長さの分布によって確認された。さらに提案法は正聴率を向上させたことから、その有効性が確認された。この方法は、他の声調言語へも適用可能である。

以上のことから本研究は、情報工学的および音声情報学的に重要な知見を与えているといえる。また最終試験の結果、上記の学生は専門的分野および関連分野の十分な知識を有することが判明した。以上のことより、学位論文審査員一同は、Win Kyawt Yinさんを博士(工学)の学位を授与するにふさわしい者とし、学位論文および最終試験をそれぞれ合格と認める。