審 査 の 結 果 の 要 旨

氏　　名　　趙 禕（チョウ　イ）

本論文は「Improving quality and flexibility of deep neural network based text-to-speech synthesis(深層ニューラルネットワーク型テキスト音声合成における自然性と柔軟性の改良)」と題し，７章より成る。近年，テキスト音声合成の技術的枠組みが，隠れマルコフモデル（HMM, Hidden Markov Model）を用いて音声の生成過程を近似する枠組みから, 深層ニューラルネットワーク（DNN, Deep Neural Network）を用いた枠組みへと大きく変遷してきた。DNN の場合，各層がどのような機能を持つのかが把握しづらいため，特定の目的のためにネットワークを修正することが難しい，という特性を持っていた。本論文では，DNN の構造に機能を与える形で，合成音声の品質と話者性に関する柔軟性の向上を検討した。比較実験の結果，従来研究を上回る性能が確認され，本研究の有効性が示された。

第一章は「Introduction」であり，本論文の背景，目的，及び，構成を述べている。

第二章は「Neural Networks for Speech Synthesis」と題し，基本的な DNN 構成である，FF 型，RNN 型，LSTM 型のネットワークを説明し，これらがテキスト音声変換にどのように応用されてきたのか，更には，音声特徴を出力するのではなく，波形値を出力するネットワークについても，説明している。合成音声は聴取実験によって評価されるが，聴取を伴わない評価（客観的評価）も可能である。客観的評価の指標についても説明している。

第三章は「BLSTM-RNN-based Multiple Speaker Speech Synthesis and Adaptation」と題し，多数話者の音声コーパスを使って，これら多数話者の音声合成を単独の枠組みで実現する，複数話者音声合成システムを深層学習に基づいて検討している。テキストから音声への変換であるが，話者非依存な変換（標準声への変換）と，後段として話者依存の変換に分け，特に後者において，話者性を付与する話者適応機能を持たせている。実験の結果，従来の多数話者音声合成システムよりも品質の良い合成音声を得ることができた。

第四章は「Speaker Representations for Multi-speaker Synthesis and Adaptation」と題し，多数話者音声合成における話者の表現方法に関して実験的な検討を行なっている。有限数の話者の一人であることを示すのであれば，one-hot vector を使えば良いが，話者集合のサイズが固定され，また，新話者の登録が原理上できない。これを解決するため，話者認識技術で標準技術となった i-vector，及び，話者モデルとして MFCC に基づく GMM を構成し，その事後確率でもって任意の話者を表現する方式を検討した。その結果，後者の話者表象が提案する複数話者音声合成の品質を向上させることが示された。

第五章は「Wasserstein GAN and Waveform Loss-based Acoustic Model Training Using a WaveNet Vocoder」と題し，生成器と識別器を対立的に学習することで，最終的に生成器の最適化を図る GAN (Generative Adversarial Netowrk)と，波高値を直接出力する WaveNet ボコーダを直接統合する枠組みとを，多数話者音声合成に適用し，その効果を検証している。実験の結果，品質向上には貢献したが，品質向上の度合いが話者に依存する様子も観測された。

第六章は「Prosody Prediction in Text to Speech Synthesis」と題し，音声合成の前処理である，テキストレベルの韻律予測において，深層学習型の予測を応用し，その精度向上を実現している。ここでは，入力情報をどの単位で扱うべきかを検討しており，統語解析エラーを回避できる文字単位の処理でも高い精度を示すことが実験的に示された。

第七章は「結論」であり，本研究のまとめと今後の展望を述べている。

以上要するに本論文は，深層学習型テキスト音声合成において，主として多数話者音声合成タスクを対象とし，1)ネットワークの機能的なモジュール化，2)話者の表現方法，更には，3)識別器と生成器の効果的な統合による話者表出の高精度化を提案し，実験的にその効果を示している。更には，4)テキストレベルの韻律予測についても頑健性の高い手法を提案し，その効果を示しており，音声工学及び情報工学に貢献するところが少なくない。
よって本論文は博士（工学）の学位請求論文として合格と認められる。