

審査の結果の要旨

論文提出者氏名 金 相勳 (KIM, Sanghun)

本論文は「Korean Corpus-based Text-to-Speech Synthesis System (大容量のデータベース基盤の韓国語の音声合成器)」と題し、トピックに対応した音声コーパスを利用した波形選択合成によって自然な韓国語音声合成を達成したものであって、全8章からなり、英文で記述されている。

第1章は「Introduction」であって、まず、音声合成の諸方式の問題点を概観した上で、それを、本論文ではコーパスベース音声合成で解決するとしている。次に、既存のコーパスベース音声合成手法・システムを整理し、本論文では、音声コーパスの効率的な構築と韻律制御について扱うとしている。また、本論文のコーパスベース韓国語テキスト音声合成 (TTS, Text-to-Speech) システムの構成図を示した上で、その処理の各過程で行った研究のポイントを整理している。最後に、第2章以降の論文の構成を述べている。

第2章は「Speech synthesis」と題し、まず、波形接続合成を中心に音声合成を概観した上で、本論文のもととなる筆者らが開発した韓国語テキスト音声合成について述べ、特に TD-PSOLA (Time Domain - Pitch Synchronous Overlap Add) によるピッチ変換の手法を説明している。次に、TD-PSOLA による合成音質の低下に言及し、本論文では、基本的にピッチ変換を行わないコーパスベース音声合成を実現するとしている。

第3章は「Text and linguistic processing」と題し、まず、韓国語の音韻体系を説明している。次に、テキスト音声合成で必要となる音韻処理について整理している。

第4章は「Corpus-based synthesis」と題し、まず、音声コーパスをトピック毎に用意する手法を提案するとともに音韻バランスを考慮して文を選択する手法を構築している。これによって、より合成環境に適合した音声素片を小さなコーパスで得られるとしている。次に、音声収録のやり方を説明し、音声認識技術を用いた収録音声の自動音素ラベリングとその精度について言及している。その上で、本論文で行ったラベリングの手修正について説明している。さらに、韻律的特徴の抽出について述べ、基本的な合成単位である tri-phone (前後の音素環境を考慮した音素) をフレーズ境界の特徴によって整理することを行っている。また、音声データベースの設計と合成単位の動的選択、データベースにない合成単位の代用規範について説明している。最後に、合成音声の評価を行って手法の妥当性を検証している。

第5章は「Prosodic phrasing」と題して、まず、本論文での合成の基本区間となる Prosodic Phrase (韻律句) について述べた上で、その境界の分類について新しい提案を行っている。韻律的特徴を用いた回帰木 (CART, Classification and Regression Tree) によってその自動抽出を行う手法を提案し、その性能を実験によって調べている。次に、テキスト音声合成で必要となる、テキストからの Prosodic Phrase 境界推定に着いて、品詞系列に着目する手法を提案し、その有効性を示している。

第6章は「Prosodic annotation」と題し、当初、英語を対象として開発された ToBI (Tone and Break and Indices) を韓国語に適応させた Korean-ToBI (前章の提案によって修正したもの) を説明し、特に Boundary Tone の自動分類について考察している。イントネーションの変化パターンを類型化して表現し、それから得られるパラメータを用いることで82%の精度で Boundary Tone の自動分類ができたとしている。

第7章は「Pruning unit instances」と題し、合成での使用頻度を重みとした重みつき VQ (Vector Quantization) で合成用音声コーパスからデータを削減する手法を開発している。聴取実験によって、音声合成品質の劣化がほとんどなく30%程度の削減が可能としている。

第8章は「Conclusions」であって、本研究で得られた成果を要約し、将来の課題について述べている。

以上を要するに、本論文は、韓国語を対象としたコーパスベーステキスト音声合成について、トピックに対応した合成用音声データベースの構築、韓国語に適した韻律ラベルの提案と自動抽出手法の開発、有効な音声データベースのプルーニング手法の開発等を行ったものである。本論文で開発された韓国語コーパスベース音声合成システムは、最高水準の合成品質を提供するものとして既に評価を受けている。コーパス音声合成の発展に大きく寄与したものであり、電子工学、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。