

本論文は「Data-Driven Generation of Fundamental Frequency Contours from Text (データベースに基づくテキストからの基本周波数パターンの生成)」と題し、音声合成において重要な課題である韻律制御について、テキストを入力とし、韻律ラベリングされた音声データベースから統計的手法により自然な基本周波数パターンを自動生成する手法を構築したものであって、全10章からなり、英文で記述されている。基本周波数パターンの生成のみならず、従来から問題とされている韻律ラベリングの手法についても研究対象としている。

第1章は「Introduction」であって、まず、高品質テキスト音声合成における韻律の重要性を指摘した上で、従来のルールベース方式とデータ主導方式の特徴、問題を明確にし、本論文で提案する手法の位置付けを明確にしている。さらに、第2章以降の論文の構成を述べている。

第2章は「A Linguistic and Prosodic Database Containing Parameters of a Superpositional Command-Response  $F_0$  model」と題し、まず、研究遂行上、必須となる韻律ラベリングされた音声データベースについて、従来のラベリングの問題点を整理した上で、基本周波数パターンの重畳モデル( $F_0$ モデル)に基づく韻律の新しい階層的記述法を提案している。

第3章は「Automatic Labeling of a Continuous Speech Database with  $F_0$  Model Parameters」と題して、第2章で提案したラベリングを自動的に行う手法を開発している。従来の $F_0$ パターンの起伏に着目する手法の限界を指摘し、データの言語情報を利用し、ラベリングに際して制約を付加する手法を開発している。

第4章は「Automatic Labeling of Accent Sandhi Types of Compound Nouns」と題して、音声合成の韻律制御で問題となる連続音声におけるアクセントサンディ現象にデータ主導の手法で対応するためのアクセントサンディデータベースの自動作成手法を開発している。

第5章は「A Linguistic and Prosodic Database with Timing Parameters Derived from a J-ToBI database」と題して、従来のJ-ToBIにより記述された韻律ラベルをもとに、A-b-S分析を併用して、第2章で提案した $F_0$ モデルに基づく韻律ラベリングを行う手法を開発し、それによって得たデータを第6章以降での韻律合成に利用するとしている。

第6章は「Data-Driven Intonation Modeling Using the Parameters of the  $F_0$  Model」であって、テキストのアクセント句についての情報をニューラルネットワークの入力とし、 $F_0$ モデルのパラメータを推定して出力する手法を開発している。 $F_0$ モデルに基づく制約により、少量のデータベースを用いたニューラルネットワークの学習で、 $F_0$ パターンの高精度の生成を達成している。

第7章は「Improvements of the Basic Configuration」であって、第6章での手法において、入力パラメータの統合を行い、性能の向上を達成している。

第8章は「Intonation Modeling Based on Binary Tree Regression」であって、回帰2分木を用いた学習により $F_0$ モデルパラメータを自動生成しており、その結果をニューラルネットワークによる手法に反映することでさらなる性能向上が図れるとしている。

第9章は「Modeling and Generation of Accent Phrase  $F_0$  Contours Based on Hidden Markov Models Synchronized at Mora Units」であって、モーラ単位で $F_0$ パターンを表現するHMMを用いて $F_0$ パターンを生成する手法を開発している。

第10章は「Conclusion」であって、本研究で得られた成果を要約し、特にニューラルネットに基づく手法の有効性に言及した上で、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、合成音声の品質を決める重要な要因である韻律の制御について、韻律ラベリングされた音声データベースを用いた学習により、テキストを入力として基本周波数パターンを自動生成する手法を開発したものであり、データ主導の手法において、モデルから得られる制約を取り入れることによって、自然な合成音声の生成を達成している。必要となる韻律ラベリングの方法、自動化等についても言及されており、音声合成分野の発展に大きく寄与するものであり、電子情報工学に貢献するところか少なくない。

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。