

審査の結果の要旨

氏名 成澤 修一

本論文は「音声の基本周波数パターン生成過程モデルの特徴パラメータ自動抽出手法とそれを用いたコーパスベース韻律生成」と題する。音声の韻律的特徴の一つである基本周波数パターンのモデルとして、その生理学的な生成過程に基づくモデルが提案されているが、本研究では、そのモデルパラメータ抽出の高精度化と、高精度化された抽出法を用いてコーパスベースの韻律生成について検討しており、全7章より成る。

第一章は「序論」であり、本研究の背景、及び目的を述べている。基本周波数パターンは音声の韻律的特徴の一つであるが、ここでは、音声によって伝搬される種々の情報のうち韻律的特徴に対応する情報を解説している。その中でも基本周波数の果たす役割が大きいこと、及び、その記述方法として生理学的知見に基づくモデルが提案されているものの、そのパラメータ抽出が困難であることを示している。

第二章は「音声の韻律分析に関する諸研究」と題し、基本周波数パターンの記述モデルとして本研究で対象にする「生成過程モデル」と、近年米国で提案され、日本語への適用も既に行なわれている「ToBI (Tones and Break Index)」について詳細に解説している。と同時に、生成過程モデルにおいて従来行なわれてきた種々のパラメータ抽出法の記述とともに各方法が持つ問題点についても言及している。

第三章は「生成過程モデルパラメータの自動抽出法」と題し、パラメータ抽出を困難にしている諸現象について列挙すると共に、その解決法の提案及び実装方式について考察している。基本周波数はその抽出時にエラーを伴うため、そのエラー除去と共に、無声区間に対する外挿、更には平滑化などの前処理を経てパラメータ抽出処理が行なわれる。基本周波数パターンの変曲点が生成過程モデルのパラメータの一つであるアクセント指令位置とおよそ一致するとの知見より、平滑化を3次曲線として近似し、その導関数の極値として指令位置を求めているなど、数学的に非常にシンプルかつ、妥当な解決方法を提案している。

第四章は「言語情報を利用したパラメータ自動抽出法」と題し、音声合成用の韻律データベース作成を念頭に置き、入力音声の書き起こしが既知である場合のパラメータ抽出精度の向上を種々の言語情報を併用することで実現している。ここでは、各種言語情報とパラメータ値との関係（規則）を実験的に構築し、精度向上を検討した。

第五章は「生成過程モデルの指令とX-JToBIラベルとの対応」と題し、同一の音声サンプルに対する生成過程モデルパラメータとX-JToBIパラメータとの比較を行なっている。前者は物理パターンのみの記述を、後者は聴覚ならびに知覚上の印象までも含めたラベリングとなっている等の相違点もあるが、前者が使用する少数のパラメータで、後者の種々のラベルを置き換えることができることを実験的に示している。

第六章は「韻律情報のデータベース化および音声合成技術への応用」と題し、本パラメータ抽出法を用いて音声コーパス（音声とその書き起こしのみのデータベース）から、韻律情報、言語情報を自動抽出し、それに基づいて音声合成時に必要な基本周波数制御規則の自動生成を試みている。生成された基本周波数パターンに対して分析再合成及び、HMM合成を用いて合成音声を作成し、聴取実験によりその自然性を評価した。韻律生成に対して全くの自動処理を行なっているにも拘らず、非常に破綻の少ない韻律生成が実装されていることを確認した。

以上を要するに、本論文は、音声情報処理の中において非常に困難な対象とされる基本周波数パターンに対する生成過程モデルパラメータ自動抽出、及びテキストからの基本周波数パターン自動生成に対してその高精度化を実現したものであり、電子情報学に貢献するところが少なくない。よって、本論文は博士（情報理工学）の学位請求論文として合格と認められる。