

平成28年度 情報工学コース卒業研究報告要旨

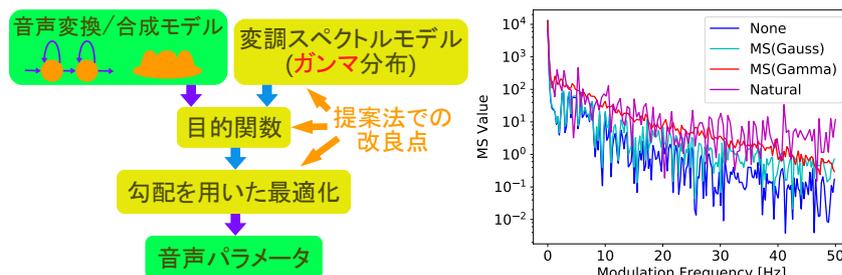
戸田 研究室	氏 名	うち の たつ のり 内 野 達 貴
卒業研究題目	ガンマ分布による変調スペクトルモデリングを用いた音声パラメータ生成法	

音声による情報伝達は、日々の生活に欠かせないものであり、所望の情報伝達を可能にする音声を生成する音声合成技術は、様々な場面で利用されている。音声合成技術の進化に伴い応用場面の拡張が期待されており、それによってより自然で様々な情報の伝達に適した音声合成技術の必要性が高まっている。音声合成技術にはいくつかあるが、その中でも音声が持つ様々な特徴を柔軟に制御できる技術として音声データから学習された統計モデルを用いて音声信号を生成する統計的パラメトリック音声合成方式が世界的に注目を集めており、さらなる発展が期待されている。

しかし、統計的パラメトリック音声合成には、合成音声の品質が、自然音声と比べると著しく劣化してしまうという本質的な課題がある。その原因の1つとして、確率モデルから音声パラメータを生成する処理で、自然音声に含まれる揺らぎ成分が統計的に誤差とみなされ、音声パラメータが過剰に平滑化されることが挙げられる。そのため、解決策の1つとして、変調スペクトルという音声パラメータの平滑化を説明するための特徴量、及びこれを考慮した音声パラメータ生成法が提案された。この手法により、過剰な平滑化が緩和され、合成音声の品質が改善することが実験的に確認されているが、その理論的な分析や考察が十分になされていない。すなわち、変調スペクトルが非負の実数値を取る確率変数なのにも関わらず、それをモデル化するために用いた確率分布が、全ての実数値を確率変数とする正規分布を用いており、適切なモデリングがなされているとは言い難い。

そこで本研究では、より理論的な見通しのよい枠組みの実現を目指し、変調スペクトルのモデリングを行う確率分布として、変調スペクトルの各成分と同じく非負の実数のみを確率変数として取るガンマ分布を利用することを提案する。ガンマ分布を導入した目的関数において、音声パラメータ系列に関する勾配を導出することで、最急降下法による音声パラメータ生成処理を実現できる。

提案法の有効性を調査するために、統計的パラメトリック音声号方式の1つである混合正規分布に基づく話者変換法を対象として、実験的評価を行った。実験結果から、①提案法において、導出した勾配を用いることで、目的関数を最大化する局所解が得られること、②従来法である正規分布によって変調スペクトルをモデリングした音声パラメータ生成法において、提案法ではより適切な変調スペクトルを持つ音声パラメータを生成できることがわかった(右図)。これらの結果から、提案法である変調スペクトルのモデリングにガンマ分布を導入した音声パラメータ生成の有効性を確認した。



(左図) 統計的パラメトリック音声合成の処理図と本研究での提案内容

(右図) 変換先音声(紫)及び生成音声(それ以外)の変調スペクトル