

音声の音響的普遍構造を用いたパラ言語情報の推定（人を観る）

研究担当者：峯松 信明（東京大学大学院・新領域創成科学研究科）

広瀬 啓吉（東京大学大学院・情報理工学系研究科）

研究代表者：峯松 信明（東京大学大学院・新領域創成科学研究科）【公募研究】

研究期間：平成 15 年度～平成 16 年度

研究 成 果 概 要

音声の音響的特徴には、性別、年齢といった非言語的特徴が不可避免的に混入する。音声から言語情報やパラ言語情報を抽出する場合、これらの特徴はノイズ以外の何者でもない。静的な非言語的特徴を表現する次元そのものを含まない音声の物理的表象は存在しないのだろうか？本研究では、この問いを、構造音韻論と呼ばれる言語学の一分野を数学的、物理学的に解釈、実装することで導出する。「構造不変の定理」と呼ばれる数学定理を導出し、その定理を音声物理に適用する形で所望の音声表象が得られる。音響イベントを音響空間内の点ではなく、分布として記述し、全ての二事象間距離を音響空間を歪める形で計算する。事象間距離を全て求めるということは、その事象群から成る幾何学的構造を規定するが、この構造は話者によって不変であるとの性質を持つことになる。この幾何学構造であるが、例えば人間の成長に伴う声道長の違いは、構造の回転として観測され、マイクなど音響機器の特性の違いは構造のシフトとして観測され、これらによって一切歪まない。しかし、発声者の言語学的、パラ言語学的側面には容易に影響を受ける。方言、訛、意図、感情などといった要因である。既に提案手法は、日本人英語音声の分析や、英語の方言識別への応用が試みられている。本研究では、パラ言語情報の推定に「音響的普遍構造」分析を用いた。

英語の母音が弱勢になると、弱母音である schwa 音に近づくことは広く知られた事実である。さて、schwa 音は母音チャートの中では、全母音の中心に位置し、これは音響的にも全母音の中心に位置することを意味する。弱勢となった母音がこの schwa に近づく、ということは、強勢から弱勢への変化は構造のサイズが小さくなる事象として解釈できる。schwa とは最もエネルギーを消費しない母音であることを考慮すると、構造サイズは調音努力の大きさを示すことが示唆されるが、英語音声を用いた強勢母音構造と弱勢母音構造の分析より、構造サイズが調音努力の差異を表すことが実験的に示された。

次に、様々な発話スタイルで「あいうえお」を発声させ、構造の部分的歪みについて検討した。演劇経験者を用いて、30 種類ほどの発話スタイルで発声させた。収録した音声に対して、任意の二発声間の差異を聴取実験によって主観評価させた。30 種類の「あいうえお」の任意の二発声間距離を求めることで、これらをボトムアップクラスタリングすることが出来る。最終的に、30 種類を 16 種類に分類し、これらの種類を自動同定する技術の構築を行なった。従来パラ言語情報の推定にはピッチ、パワー、継続長といった韻律的特徴が用いられる。本研究で提案している普遍構造分析は、音声の分節的特徴のみを用いて、音声の超分節的特徴を定義していることに相当し、音声を韻律的特徴と分節的特徴とに分離することを前提とした方法論では定義不可能な特徴量となっている。上記したように調音努力は構造サイズに影響を与える。これは、継続長さ、パワーの情報（の一部）が既に構造の中に埋め込まれていることを意味する。ここでは、基本周波数の情報と構造分析との交互作用について検討した。実験の結果「韻律的特徴及びそれを用いた機械学習」という従来方法よりも、「普遍構造解析+基本周波数パターン」という提案手法が、より頑健かつ高精度にパラ言語情報の推定が可能であることを実験的に確認した。本構造解析は、音声認識（即ち言語情報の抽出）も可能であることが示されており、提案手法は、言語情報とパラ言語情報を同時に推定可能な枠組みであると言える。音声を分節と韻律に分け（要素還元）、前者から言語情報を、後者からパラ言語情報を抽出していた従来法と比較して、本手法は音声の全体的表象を定義し、それに基づいて処理系を構築している。

論 文 発 表 等

1. N. Minematsu, S. Asakawa, and K. Hirose, "The acoustic universal structure in speech and its correlation to para-linguistic information in speech," Proc. Int. Workshop on Man-Machine Symbiotic Systems (IWMMS'2004), pp.69-79 (2004)
2. 浜野紘一、峯松信明、広瀬啓吉、"音声の分節的特徴に着眼したパラ・非言語情報推定に関する実験的検討", 電子情報通信学会音声研究会, SP2003-197, pp.25-30 (2004)
3. 朝川智、峯松信明、広瀬啓吉、"音響的普遍構造のサイズと単語境界における音響的分離度に着眼した米語音声の音響分析", 電子情報通信学会音声研究会, SP2004-28, pp.53-58 (2004)
4. 藤野真紀、峯松信明、広瀬啓吉、"音声の音響的普遍構造に着眼したパラ・非言語情報推定に関する実験的検討", 日本音響学会春季講演論文集, 2-5-3, pp.59-60 (2005)