

平成26年度 情報工学コース卒業研究報告要旨

大西 研究室	氏 名	池 田 直 史
卒業研究題目	音源分離を用いた発話重複区間検出	

複数話者による談話では、話者が同時に発言することによる発話の重複がしばしば発生する。このような発話の重複は、談話分析作業の妨げや、音声認識精度の低下などにつながる。同時発話を音声認識するためには音源分離が必要となるが、そのための前処理として発話が重複している区間を検出することが非常に重要となる。

発話重複区間検出の従来手法として、HMMを用いた手法による検出があるが、HMMネットワーク構成のために事前学習が必要であることや、検出精度があまり高くないことが問題点として挙げられる。

そこで本研究では、音源分離を用い、事前学習が不要な発話重複区間検出の手法を提案する。提案する手法では、話者数分のマイクを用いて収録した観測信号に対して、FDICA(周波数領域独立成分分析)による音源分離を用いることで発話重複区間検出を行う。統計的に独立な重複音声の分離音は互いに無相関な信号となる。しかし、単独発話について音源分離を行うことは難しく、分離音同士の相関は重複音声の分離音同士と比べて高くなる。提案手法は、この相関の違いを利用して発話重複区間の検出を行う。相関は、時間周波数領域での時間方向50点(0.8s)の相関係数として、1点ずつシフトしながら求める。これにより得られた時間方向の相関係数から、図1のように、相関係数が閾値以下となっている部分を重複発話区間と判断する。

提案手法と従来手法(HMMを用いた手法)を用いて、重複区間検出実験を行った。使用した音声は、2人の話者による重複を含む発話で、計100セット200発話である。提案手法を用いた実験は、マイク間の距離(5cmと80cm)、マイクの指向性の有無を変えて4通りの条件で行った。時間周波数領域の1点(0.016s)毎に重複発話であるか否かを判定し、検出精度はF値により求める。その結果、適切な閾値を与えれば、どの場合についても7~8割程度の検出精度が得られ、従来手法(検出精度は5割程度)よりも良い結果となった。しかし、マイク間の距離が離れている場合には、音源分離をする前の観測信号同士についても重複区間の相関係数が低い値となっており、音源分離をしなくとも重複区間の検出ができると考えられる。以上のことから、提案する手法はマイクロホンアレイを使用するなどマイク間の距離が近い状況で複数人の話者による発話を収録する場合において有効であると言える。

今後の課題としては、検出の際に用いる閾値を、音源分離の結果から自動で決定できるようにすることが考えられる。

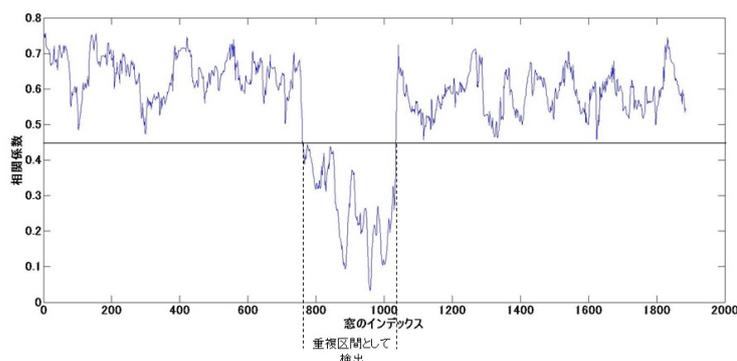


図1. 時間方向の相関係数からの重複区間検出