論 文 の 内 容 の 要 旨

論文題目　Study on Efficient Prior Control for Realizing Practical Systems of Speech Recognition
（音声認識の実用化のための効率的な事前制御技術に関する研究）

氏　名　小橋川　哲

近年、高性能のクラウド上の計算機とスマートフォンの普及等により音声検索等の音声認識応用アプリケーションが一般に使われ始めている．しかしながら，背景雑音が混入する環境や曖昧な発声を含む話し言葉音声に対しては，ベースとなる音声認識精度は大幅に劣化するため，実用化のための障壁となっている．そこで，本研究では，実用化のために生じる課題への対応するため，音声認識の利用場面に応じて活用できる前提条件に基づく事前制御技術を導入する事で，通常の音声認識処理では得られない効果を得る事を目的とする．

本研究では，a)タブレット端末利用時における耐雑音音声認識，b)コンタクトセンタや議会における話し言葉音声認識，の2つの大きな研究課題を対象とする．音声認識の利用場面としては，i)タブレット端末上での音声インタフェース，ii)コンタクトセンタ音声マイニング，iii)議会音声書き起こし支援，の3つを想定する．ここで，タブレット端末上の音声インタフェースに対しては，高いレスポンスのもとで，雑音耐性の強化が必要になる．また，コンタクトセンタ音声マイニングに対しては，大量の蓄積通話音声データから高い認識精度な音声ドキュメントを収集する事が求められる．議会音声書き起こし支援に対しては，限られた処理時間制約のもので変動する話者・環境に対して高い認識精度が求められる．

まず，タブレット端末利用時では，S/N比の変動および空間伝達特性の影響に対応するため，
「1) 事前観測雑音を用いたモデル適応と正規化」技術を開発した．本技術は，定常的な背景雑音は事前に観測が可能という前提を置き，事前観測した加法性雑音のみを用いて，スペクトルサブトラクション(SS: Spectral Subtraction)による加法性雑音抑圧手法，モデル合成法(PMC: Parallel Model Combination)による加法性雑音適応手法，ケプストラム正規化(CMN: Cepstral Mean Normalization)による乗法性雑音正規化手法を融合させて，S/Nの変動や空間伝達特性である乗法性雑音(歪み)に強い音声認識方式である．

次に，コンタクトセンタでは，限られた処理時間要件の下で精度向上に貢献する「2)フレーム独立信頼度を用いた事前高速教師なし適応」技術を開発した．本技術は，低遅延のリアルタイム処理は必須ではない蓄積された音声という前提を置き，対象音声データを音響モデル中の限られたGMM(Gaussian Mixture Model)を用いる事により，最尤線形回帰法(MLLR: Maximum Likelihood Linear Regression)と性別選択を融合させた高速に精度向上ができる音声認識方式である．

また，コンタクトセンタでは，信頼度が低く後段のテキストマイニング処理に悪影響を及ぼすため棄却される音声データの音声認識処理に過剰な計算コストを大幅に削減する「3)事前信頼度によるデータ選択」技術を開発した．本技術は，大量音声データが蓄積されている前提を置き，低精度の音声認識結果は不要であることから，音響モデル中の限られたGMMを用いて高速に信頼度を推定して，精度が高い音声のみを音声認識対象とする方式である．

さらに，コンタクトセンタ向けに認識精度が低く後段で棄却すべき低音質の音声データに対する音声認識処理の計算コストを削減する「4)事前ビーム幅制御による認識処理時間安定化」を開発した．本技術は，大量音声データの中に含まれる低精度・低品質音声には計算量が無駄にかけなくて良いという前提を置いて，対象音声データを音響モデル中の限られたGMMを用いて予め高速に走査して得られた事前スコアから，音声認識処理中の探索空間の広がりを制御する事で，音声認識処理時間を安定化させる方式である．

加えて，議会録作成支援のための音声認識システムの実用化に向けて，議会場の収音環境および話者・環境の変動に対応するための「5)議会録作成支援のための高速事前音響処理」を開発した．本技術は，対象の議会音声が区分化された単位で蓄積されて送られてくる前提を活かし，チャネル選択，話者インデクシング，特徴量正規化，教師なし適応といった複数の事前音響処理を高速に行う事で，限られた計算時間の下で，変

動する音声に対して高い音声認識精度を実現する方式である.

以上，5つの方式はいずれも音声認識の利用場面に即して活用できる前提を置くことで導入可能な事前制御に関する処理であり，従来の音声認識方式では得られない効果を得る方式である.