



誤りから学ぶ音声認識システム

～音声認識のための識別モデル学習法～

Speech Recognition System that Learns from Errors

- Discriminative Model Training for Speech Recognition -

概要：音声認識システムは音響モデルと言語モデルを利用して音声をテキストに変換します。高い認識率を得るにはモデルの学習技術の向上がとても重要です。人間の勉強法のひとつに問題の解答と答え合わせを通じた学習がありますが、我々は同じ仕組みをモデル学習にも導入し、高精度な音声認識を実現しています。

① 音声を聞き



Q
出題



A_{Q,Λ}
解答

② 認識してみる



多くの場合あーのー聴覚的フィードバックつまりいーシステムかわ返ってくる音声というのはあーま母語話者のターゲットとなる温泉が殆どなんだと思うんですけども

テストの復習のように、自分の解答と答えを照らし合わせて学習することで、更に賢い音声認識を実現します。

学習

採点

$$\Lambda \leftarrow \arg \max_{\Lambda} O_{\Lambda}(A_{Q,\Lambda})$$

$$O_{\Lambda}(A_{Q,\Lambda})$$

多くの場合あーのー聴覚的フィードバックつまりいーシステムかわ返ってくる音声というのはあーま母語話者のターゲットとなる温泉が殆どなんだと思うんですけども

から

音声

脱落

④ 同じ間違いをしないように学習

③ 答えを見て



連絡先 (Contact information)

大庭 隆伸 (Takanobu OBA)

信号処理研究グループ (Signal Processing Research Group)