

ISTCにおける 音声認識ソフトウェアの開発

河原達也
(京都大学)

これまでの経過

- 1995～1997 IPSJ/SLP傘下 WG
 - JNASコーパスの設計
- 1997～2000 IPAプロジェクト
 - 日本語ディクテーション基本ソフトウェアの開発
- 2000～2003 連続音声認識コンソーシアム CSRC
 - 認識ソフトウェアの改善
 - 音響・言語モデル等の充実
- 2003～2008 e-Society基盤ソフトウェア
- 2003～2006 音声対話技術コンソーシアム ISTC
 - 対話システムを指向した音声認識の改善

ISTCでの主要開発目標

- 音声認識エンジンJuliusの性能改善・機能追加
- 音声認識エンジンJuliusのSAPI/SALT対応
- 音声認識ソフトウェアのカスタマイズツール
- 音声合成エンジンgtalkのSAPI対応

Juliusの最近の性能改善・機能追加

- 探索アルゴリズムの細かな改善
- 認識結果への信頼度付与
- クラスN-gram 対応
 - 非定型な文に対して頑健な認識
- HMMファイルのバイナリ化
 - 高速な立ち上がり(対話用モデルなら1秒)

JuliusのSAPI/SALT対応

- Julius for SAPI ver 2.3 リリース(02/05)
 - Julius rev.3.4対応
 - マイクロソフト SA-SDK ベータ3対応
 - SALTでの動作確認
 - いくつかのノウハウ→サンプル付属のメモ参照

新年度からの開発

- 音声認識ソフトウェアのカスタマイズツール
 - 適当な音響モデル・言語モデルの組合せ
 - Juliusの自動コンフィグレーション
 - Juliusの標準jconfファイル作成
- 音声合成エンジンgtalkのSAPI対応

オプション(要望があれば)

- 新聞言語モデルの更新(伊藤)
- スーパー音響モデル(三村)
- 対話用言語モデル
- 対話用音響モデル
- 英語用ディクテーションキット
- 話者認識ツールキット

講習会

- 音声認識
- 音声合成
- 顔出力

デモンストレーション

- Julius+gtalk+SAPI
1. ソフトウェアロボット操作
 - Julius SAPIのサンプルプログラム
 2. ソフトウェアサポートKB検索
 - SALT
 - まもなくリリース

関連Webページ

- 連続音声認識コンソーシアム(CSRC)
 - 現在もCD-ROM購入可能、最終版は2004年3月予定
 - <http://www.lang.astem.or.jp/CSRC/>
- Julius
 - 最新版(SAPI対応版含む)フリーダウンロード
 - <http://julius.sourceforge.jp/>
- Microsoft Speech Application SDK
 - SALTには必要
 - <http://www.microsoft.com/speech/>
- [e-Society基盤ソフトウェア「自然な音声対話処理」](#)