

ISTCにおける 音声認識ソフトウェアの開発状況

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(李先生)
- 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 3.4.xにおける 性能改善・機能追加(李先生)

- 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)

Rev. 3.4.2(最新版)は5月7日リリース

JuliusのSAPI/SALT対応

- Julius SAPI版 ver2.3 リリース (2月21日)
 - Julius rev.3.4対応
 - Windows XPのSAPI対応
- マイクロソフトSASDK 1.0が正式リリース (5月7日)
 - SALT含む
 - Julius SAPI版によるSALT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
- Julius SAPI版用スタンダードモデルキット、サンプルプログラム・文法などを更新 (5月13日)

今後の開発計画

- 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
- CSRC最終版にバンドル
- 来年度ISTCに含めたい

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使用

関連Webページ

- 連続音声認識コンソーシアム(CSRC)
 - 最終版を一般頒布中
 - <http://www.lang.astem.or.jp/CSRC/>
- Julius
 - 最新版(SAPI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」](#)

講習会

- 日時:8月23日(月)~27日(金)
- 場所:京都大学学術情報メディアセンター
- ISTC一般会員は無料(3名まで)
- 内容:
 - (前半)音声認識の基礎から
 - (後半)音声合成、顔画像合成、音声・MM対話
- 申込...至急(遅くとも今週中)
 - <http://www.ar.media.kyoto-u.ac.jp/seminar/onsei0408.html>