
부산대학교 Text-to-Speech 시스템 (PNU-TTS) 소개

2000. 3. 31.

김형순

부산대학교

PNU-TTS의 주요 특징

- 정현파 모델 기반의 합성방식 채택
 - ↳ 음질을 유지하면서 다양한 변환 가능
 - CART 기법을 기반으로 한 통계적 운용 모델
 - 문맥종속 초중성(CV) 및 중성(C) 단위의 음성 DB
 - ↳ 여성 16 MB 및 남성 24 MB의 저장량 DB size
 - 음색 제어 기능
 - ↳ 단일 DB로 다양한 화자의 음색 합성
 - Microsoft 사의 Speech API(SAPI) 규격 만족
 - 상업적 용도로도 무상 사용 가능 (구문분석 DLL 제외)
-

PNU-TTS 사용 환경

□ 사용환경

- ⇒ OS: MS Windows 95/98/NT 4.0 이상
- ⇒ CPU: Intel Pentium 이상 또는 호환
- ⇒ RAM: 32 MB 이상

PNU-TTS의 음성생성방식

□ 정현파 모델 기반의 합성방식

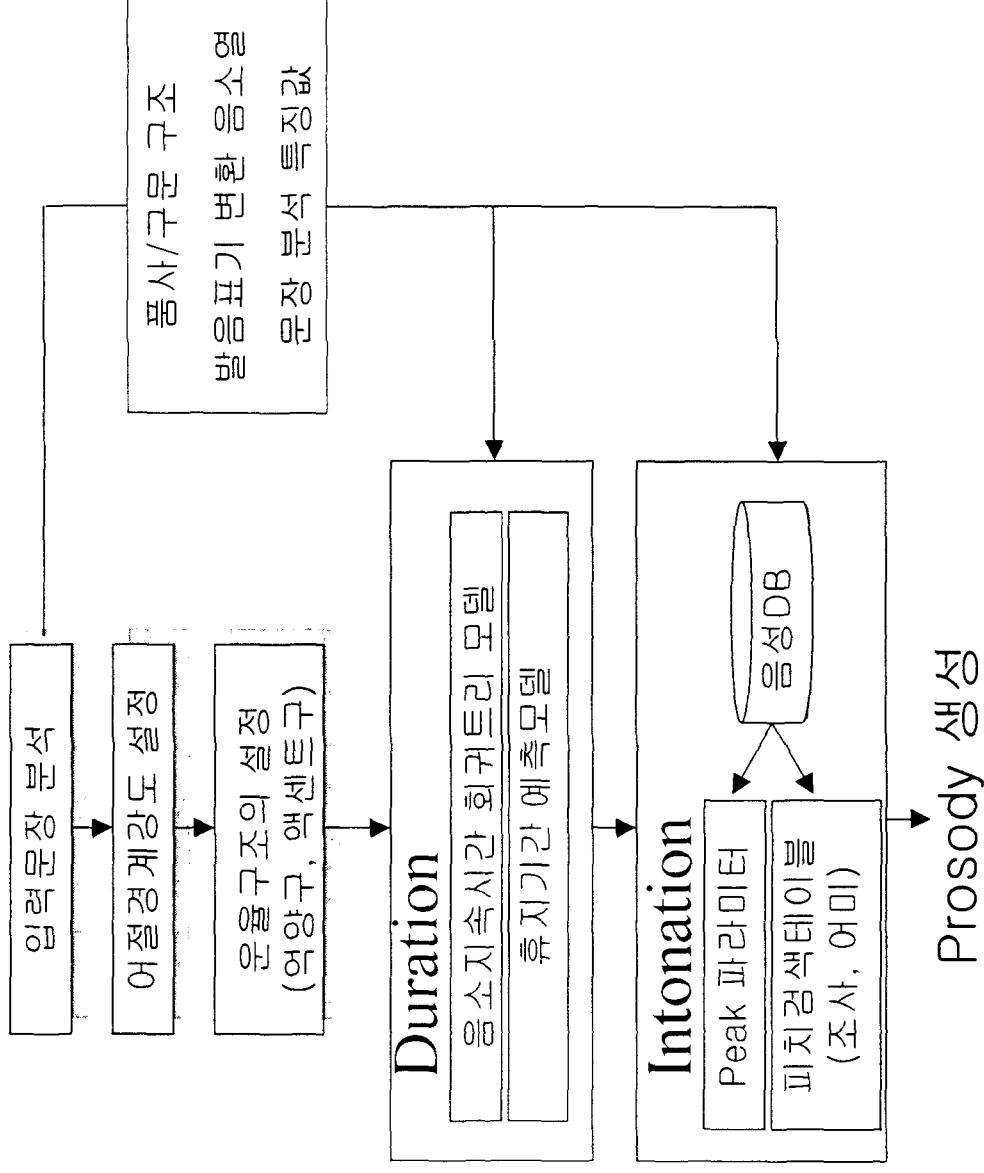
⇒ 단구간 음성신호를 정현파들의 합으로 표현

$$s(n) = \sum_l A_l \cos(\omega_l n + \theta_l)$$

⇒ 기존의 PSOLA 방식에 비해 계산량은 많으나,
음질을 유지하면서 다양한 변환이 용이함

⇒ 단일 DB로부터 다양한 음색(화자특성) 표현 가능

PNU-TTS의 운율 제어부



PNU-TTS 음성 DB 구축

- 합성 단위
 - ⇒ 좌우문맥을 고려한 초/중성(C¹V) 및 종성(C¹) 단위
 - ⇒ 총 6,080 개의 합성단위
 - ⇒ DB size : 여성 16 Mbytes, 남성 24 Mbytes 미만
 - ⇒ 추가적인 DB 용량 감축 가능
- DB 구축을 위한 음성 코퍼스
 - ⇒ 남녀 아나운서 : 각각 1,000문장 및 500단어 발음

PNU-TTS 라이브러리 (DLL) 구성

- 문장분석 DLL
 - ↳ Text 전처리, 구문분석, 발음표기변환
- 운율생성 DLL
 - ↳ 억양, 지속기간 및 휴지기간 정보 생성
- 음성합성 DLL
 - ↳ 문장분석 및 운율정보로부터 음성파형 생성
- 3개의 DLL을 함께 또는 개별적으로 사용할 수 있음
 - ↳ TTS 모듈 중 일부분만 자체 개발하여 적용 가능
(소규모 연구팀의 TTS 연구에 도움 제공)

PNU-TTS 사용 절차

- 사용 조건
 - ⇒ 연구용 및 상업적 응용을 위해 무상 사용 가능
 - ⇒ 단, 문장분석 DLL에 포함되는 형태소분석 DLL은 한성대 강승식 교수의 소유이며, 상업적 이용시 별도의 license를 구입하여야 함
- 사용 절차
 - ⇒ 부산대 음성통신실험실 홈페이지를 통해 평가판 신청
<http://voice.ee.pusan.ac.kr/speech/tts.htm>
 - ⇒ 상업적 이용시 별도의 계약 체결