FONIX

VoiceSync™ SDK

Lip Sync Applications

VoiceSync 1.0 — a significant leap forward in speech recognition technology

Lip Sync for Animated Characters

The use of avitars and animated characters has grown significantly and the demand from users for better and better animations requires more natural lip and facial movements when the characters are speaking. Fonix VoiceSync generates the phonetic data and timing to create the 'human-like' lip and facial movements.

Fonix VoiceSync 1.0 for Lip Sync Applications

VoiceSync automatic speech recognition (ASR) is based on Fonix proprietary neural network technology and Fonix VoiceIn 4.2. It provides accurate phonetic information about the speech that allows developers of animation products to more closely match lip, tongue, and facial animations to natural human movements. The Fonix VoiceSync software provides the phonemes and timing information and the corresponding markings for the audio stream. Fonix VoiceSync operates on scenariors where the text is provided as well as for applications, such as online voice chat, where only the audio is provided.

Fonix VoiceSync provides a low memory and MIPS solution that will run on your game hardware without significantly

---

Fonix VoiceSync Features

- Recognize phonemes in live or recorded speech
- Phonetic engine recognizes phonemes aligns them with their location in the speech audio.
- Two modes of operation:
  - Utilize text of the utterance to provide the most accurate phonemes and their locations
  - Generate phonetic estimates of free-form speech (e.g. online voice chat)
- Based on the Fonix VoiceIn 4.2 Speech Recognition Technology and SDK

Accurate

- Accurately identify phonemes and their locations
- Noise-tolerant
- Insensitive to environment acoustics (no training necessary)

Efficient

- Compact neural networks/small memory footprint
- Small computing requirements (starts at 20 MIPS)
- Real-time operation - 70 - 100 ms delay

Flexible

- Speaker-independent, continuous speech
- Normalized confidence scoring for phonemes, words or phrases
- Wide character support

Vocabulary

- No vocabulary limitations

Languages

- UK/US English, Canadian/European French, German, Italian, Japanese, Korean, Castilian/Latin American Spanish

Memory Usage

- Base Engine: 352 KB ROM
- Dictionary: 0 – 2,323 KB ROM
- Neural Networks: 50 – 225 KB ROM
- approx. 200 KB

Processor Utilization

- Word spotting - minimal CPU usage

Audio Requirements

- 8/16 KHz 16 bit PCM

Platforms

- Windows 32/64
- Linux 32/64
- Microsoft Xbox 360
- Sony Playstation 3
- Nintendo Wii

Developer Support Program: SDK training, application development.

fonix
freedom of speech™

Fonix Corporation
387 South 520 West STE 110
Lindon, UT 84042
801-553-6600
www.fonix.com