

# ISD-SR3000

The **ISD-SR3000** is a complete embedded speech recognition processor. It consists of a speech recognition engine, a speech compression engine, and a recording function. The **ISD-SR3000** hardware includes a parallel RISC/DSP core with an optimized instruction set, a flexible CODEC interface, and a serial host controller interface. The speech recognition engine uses sophisticated Hidden Markov Models (HMMs), which enable recognition of continuous speech and connected digits. An application consists of speaker-independent commands (chosen by the application developer), connected digits and speaker-defined commands. The speaker-defined commands allow users to store and recognize voicetags that can be used for custom commands or name list management. The speaker-defined commands use the HMMs, providing much more robust performance compared to conventional speaker-dependent commands. The speaker-independent commands, the audio prompts, and the speaker-defined voicetags are stored in external memory, allowing for maximum application flexibility. Typical storage requirements are 2kB for each pre-defined command, 2.5kB for each audio prompt, and 3kB for each voicetag, including the model, recording, and data. Application commands are divided into topics

## Ideal command and control applications

- Internet appliances
- Hands-free car kits
- Control of automotive accessories and systems
- Desktop phones
- Instrumentation control
- Remote control of consumer devices

## Speech Recognition Processor for Command and Control Applications



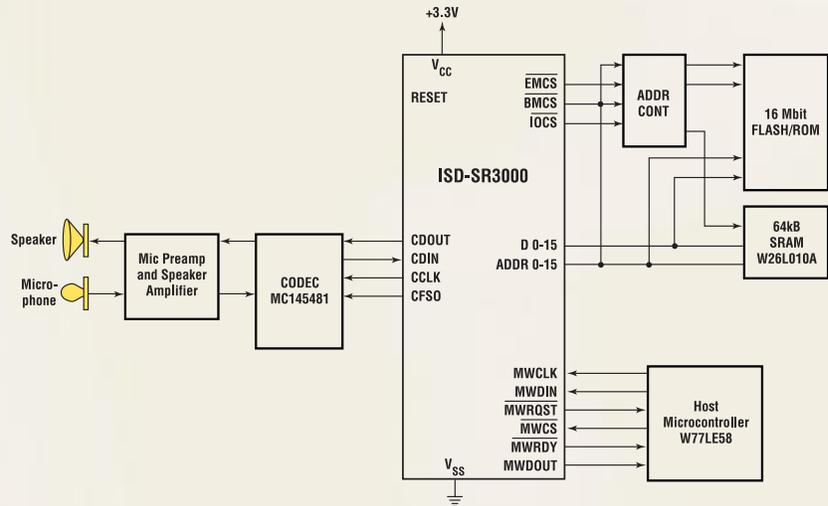
(menus), with active vocabulary size governed by the size of the external SRAM. A development system, the **ISD-DS3000**, is available. The development system includes tools for compiling and sizing commands and prompts, as well as sample C-code for host control program development.

### Speech Recognition Attributes

- Continuous recognition
  - Activated by key word or leading silence
  - Rejects invalid vocabulary
  - Users do not have to alter speech patterns
- Connected digit recognition
  - No domain restrictions
  - Returned after trailing silence
  - Continuous strings and inter-digit pauses allowed
- Speaker-independent command recognition
  - No user training required
  - Application-specific commands
  - Defined by the application developer
  - Number of commands determined by external memory size
  - Typical applications support up to 100 total words, 30 active at any time
- Speaker-defined voice tags
  - Built from HMM-based acoustic models
  - Recording, model, and data for each tag stored on external non-volatile memory
  - More robust than speaker-dependent word models
  - User enters name only one time
- American English and German languages
  - Contact Winbond about language development

## Stand-alone speech recognition system diagram

ISD-SR3000



### Recognition Technology

**ISD-SR3000** uses a statistical Hidden Markov Model (HMM) recognition engine. The sampled speech utterance is split into distinct frames. The DSP engine then analyzes the frames, extracting frequency formants that describe the speech energy information for each frame. These are further quantized into feature vectors that are processed by the RISC engine using a high-speed search algorithm. The search algorithm compares the feature vectors to the stored HMMs using statistical processing techniques. This process is ideal for recognition of continuous speech. The output of the **ISD-SR3000** is a token that corresponds to the stored word list. The token is used by a host controller in the same fashion as a keypad.

### Recognition Processor Attributes

- Combined DSP and RISC processor
- Optimized speech processing instruction set
- Interfaces to  $\mu$ -Law, A-Law, or Linear CODEC
- Serial interface to host controller
- Power supply: +3.3V or +5V
- Current: 40mA (typical) during active recognition
- Package: 100-pin QFP

### Application Attributes

- Enables hands-free control of communication, consumer, automotive, and industrial applications
- Advanced API supports recognition, recording and compression functions
- Compression functions enable audio prompts for user feedback
- Multiple topic activation provides maximum application design flexibility

### Application Development Tools

There are a number of documents contained in the Winbond-USA web site, including a complete product data sheet, a design procedure, and an application development tutorial: [www.winbond-usa.com/products/speechrecog/index.html](http://www.winbond-usa.com/products/speechrecog/index.html)

An application demonstration and development tool, the **ISD-DS3000** is available for basic technology evaluation. Contact Winbond for information about the SDK for this tool. The **ISD-DS3000** supports a sample application including the C-code for the host controller. Reference designs for car kits, remote control, and telephones will also be available. Contact our applications staff at [recho@winbond.com](mailto:recho@winbond.com) for further information.

To order products or for more information:

Winbond Electronics Corporation America  
 2727 N. First Street  
 San Jose, CA 95134  
 1-800-677-0769 (US Only), 408-943-6666  
 Fax 408-544-1789  
 e-mail: [info@winbond-usa.com](mailto:info@winbond-usa.com)  
 Web: [www.winbond-usa.com](http://www.winbond-usa.com)