The RSC-4128 represents Sensory's next generation mixed signal processor for speech and analog I/O. The RSC-4128 is designed to bring advanced speech I/O features to cost sensitive embedded and consumer products. Based on an 8-bit microcontroller, the RSC-4128 integrates speech-optimized digital and analog processing blocks into a single chip solution capable of accurate speech recognition; high quality, low data-rate compressed speech; and advanced music. Products can use one or all features in a single application.

The RSC-4128 operates in tandem with FluentChip™ firmware, an ultra compact suite of recognition and synthesis technologies. This reduced software footprint enables, for example, products with over 150 seconds of compressed speech, multiple speaker dependent and independent vocabularies, speaker verification, and all application code built into the RSC-4128 as a single chip solution. Revolutionary Text-to-Speaker-Independent (T2SI™) technology allows the creation of SI recognition sets by simply entering text.

In addition to improved recognition performance, the RSC-4128 provides further on-chip integration of features. A complete speech I/O application can be built with as few additional parts as a clock crystal, speaker, microphone, and few resistors and capacitors.

FEATURES

FULL RANGE OF FLUENTCHIP™ CAPABILITIES

‣ Enhanced Word Spotting capability (10 SI or 4 SD words) in parallel
‣ Noise robust Speaker Independent, Dependent & Continuous Listening recognition
‣ High quality, 3.7-7.8 kbps speech synthesis & sound effects with Sensory "SX" synthesis technology
‣ Speaker Verification (SVWS) – Noise robust voice biometric security
‣ 8 voice MIDI-compatible music synthesis coincident with speech; drum track feature enables additional voices
‣ Voice record & playback
‣ Audio wake up from sleep

INTEGRATED SINGLE-CHIP SOLUTION

‣ 8-bit microcontroller
‣ On-chip 16 bit ADC, 10 bit DAC and microphone pre-amp
‣ Independent, programmable Digital Filter engine
‣ Uses low cost 3.58MHz crystal (internal PLL)
‣ 4.8 KB total RAM (256Bytes user-RAM)
‣ Five timers (3 GP, 1 Watchdog, 1 Multi Tasking)
‣ Twin-DMA, Vector math accelerator, and multiplier
‣ External memory bus: 20-bit Address(1Mbyte), 8-bit Data
‣ On chip storage for SD, SV, templates (10 templates)
‣ Code security through no ROM dump capability
‣ Built-in Analog Comparator Unit (4 inputs)
‣ Low EMI design for FCC and CE requirements
‣ 24 configurable I/O lines with 10 mA (typical) outputs
‣ Fully nested interrupt structure with up to 8 sources

LONG BATTERY LIFE

‣ 2.4 – 3.6V operation
‣ 12mA (typical) operating current at 3V
‣ 2 low power modes; 1 μA typical sleep current

RSC-4128 Block Diagram
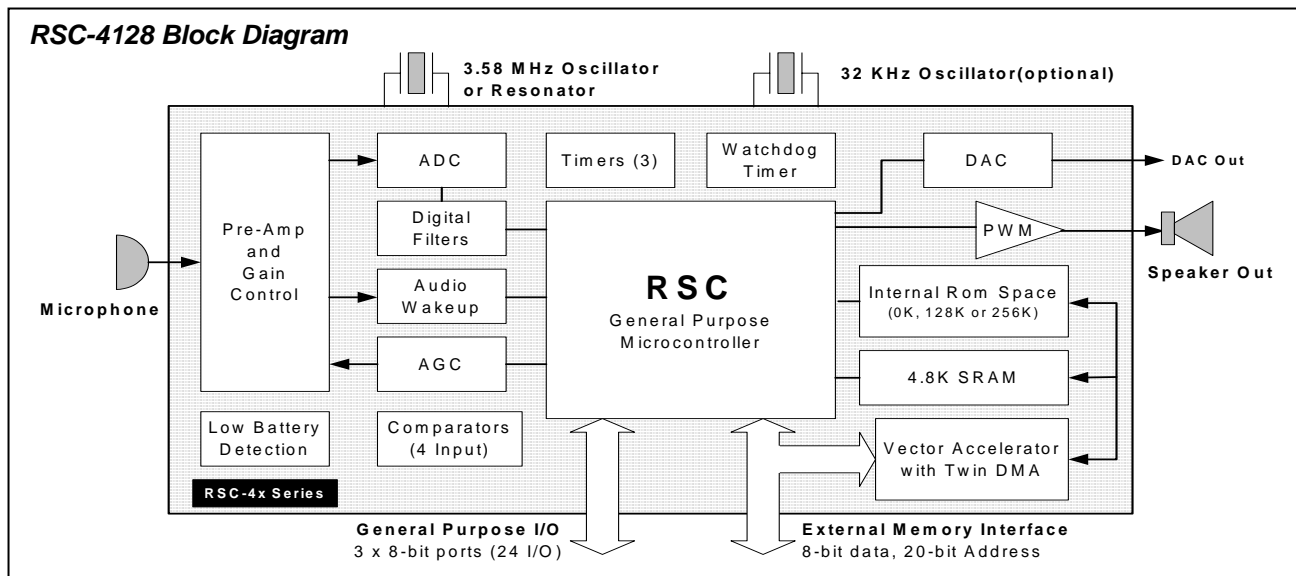
The RSC-4128 is designed to support HMM (Hidden Markov Modeling) as well as Neural Network technologies provided in FluentChip™ firmware to perform speaker independent (SI) speech recognition. This requires on-chip or off-chip ROM to store the words to be recognized. Speaker dependent (SD) recognition requires programmable memory to store personalized speech templates and may be on-chip SRAM or off-chip Serial EEPROM, Flash Memory, or SRAM. The RSC-4128 has several additional speech recognition features as described below:

- ‣ **Speaker Independent** recognition requires no user training. The RSC-4128 can recognize up to 20 words in an active set (number of sets is limited only by internal ROM or external memory size). Text-to-SI (T2SI) recognition, based on HMM technology, allows creation of SI recognition sets in seconds.
- ‣ **Speaker Dependent** recognition allows the user to create names for products or customize vocabularies. Up to 100 words can be recognized in an active set (number of sets is limited only by internal ROM or external memory size). The RSC-4128 can store up to 10 SD words in on-chip SRAM.
- ‣ **Continuous Listening** allows the chip to continuously listen for a specific trigger word. With this feature, a product "activates" when a specific word is spoken, framed by quiet before and after. Continuous listening provides the lowest false fire rate for trigger words.
- ‣ **Word Spotting** allows the chip to continuously recognize for up to 10 SI or 5 SD words at a time. In word spotting mode, the word(s) to be recognized may be spoken in the middle of speech.
- ‣ **Speaker Verification** technology allows the RSC-4128 to be able to identify whether a particular word is spoken by the original speaker. Up to 10 SV templates can be stored on-chip, or more with external programmable memory

SPEECH AND MUSIC SYNTHESIS

The RSC-4128 provides high-quality speech synthesis using state of the art "SX" technolgy. Typical data rates are 6000 bits per second and require on-chip or off-chip ROM to store audio sound data. The RSC-4128 uses a MIDI-like system to generate high-quality, eight-voice, wave table music synthesis.

RECORD AND PLAYBACK

The RSC-4128 can perform speech record and playback (sometimes called "voice memo") at various compression levels depending on the quantity and quality of playback desired. The record and playback technology also performs silence removal to improve sound quality and reduce memory requirements.