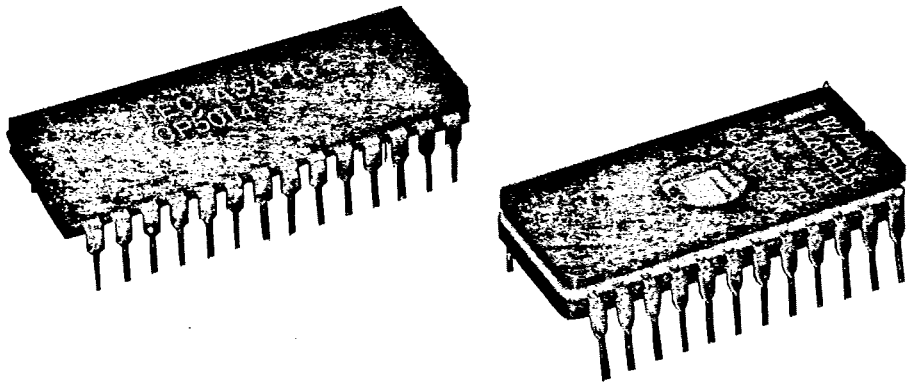
VocaLink™

INTERSTATE VOICE PRODUCTS
A FIGGIE INTERNATIONAL COMPANY

Model VRC100-2A — high accuracy in a low-cost voice recognition chip set.



Highly accurate real-time operation 98-99%+* and up to 200-word vocabulary

Trainable for any vocabulary in any language

Usable with direct microphone, wireless communication, or telephone

Selectable decision threshold for rejection of unwanted inputs

Input/output port configuration for easy product integration

17 user commands for maximum flexibility

High-Accuracy Voice Recognition in a Chip Set

Voice Recognition Chip Set VRC100-2A, figure 1, consists of two chips used as building blocks for speech recognition systems and capable of recognizing as many as 100 words or short phrases. This capability may be expanded to 200 words by use of additional RAM. The VRC100-2A permits a flexible system design to meet user requirements. The two integrated circuits are designated ASA-16 and 100-2A. Figure 1 shows a simplified block diagram using the VRC100-2A.

Speech input is analyzed by a 16-channel spectrum analyzer within the ASA-16 speech preprocessor and converted to a digital representation of the characteristics of the spoken input. This digital data is then converted to a fixed-size pattern that preserves the information content of the spoken inputs while discarding redundant features. During word training, these patterns are used to derive reference templates for each vocabulary item. The templates are then used in the recognition process for comparison with incoming spoken words. Vocabulary templates are stored in an external RAM, while the processing algorithms are contained in the recognition/control program (EPROM).

17 User Commands for Maximum Flexibility

This EPROM firmware accommodates 17 user commands for user flexibility. There are two training commands. The normal training command is issued when all or part of the specified vocabulary is cleared and then trained a selectable number of samples. The training update command is used when the stored reference patterns of the specified vocabulary are to be augmented by additional training.

The VRC100-2A training algorithm automatically rejects utterances during training that do not sufficiently agree with the same utterance from previous training samples of the word. This prevents significant alteration of a vocabulary reference pattern caused by spurious noise (bumping the microphone, door closure, coughing), speaking inconsistencies, or simply failing to utter the prompted vocabulary item. In such an event, it may be necessary to repeat an utterance before being prompted to the next sequential utterance. User prompting is accomplished via the host processor and display. Verification is provided during training by the VRC100-2A.

The reset command is used to initialize the recognizer and the associated RAM. Reset causes the word boundary and adaptive gain parameters to return to default value. The VRC100-2A chip set also allows rejection control of invalid utterances via the set reject and read reject threshold commands. Two commands allow reading or writing of word boundary and background noise parameters. Another command allows the object code to be downloaded and executed.

The major operational command of the VRC100-2A is the recognize command. This command allows recognition of any specified vocabulary up to 200 words. The command allows recognition of both contiguous and/or random syntax with one or more common subvocabularies.

*Performance of 98-99%+ recognition accuracy on test vocabularies in 100 word subsets.

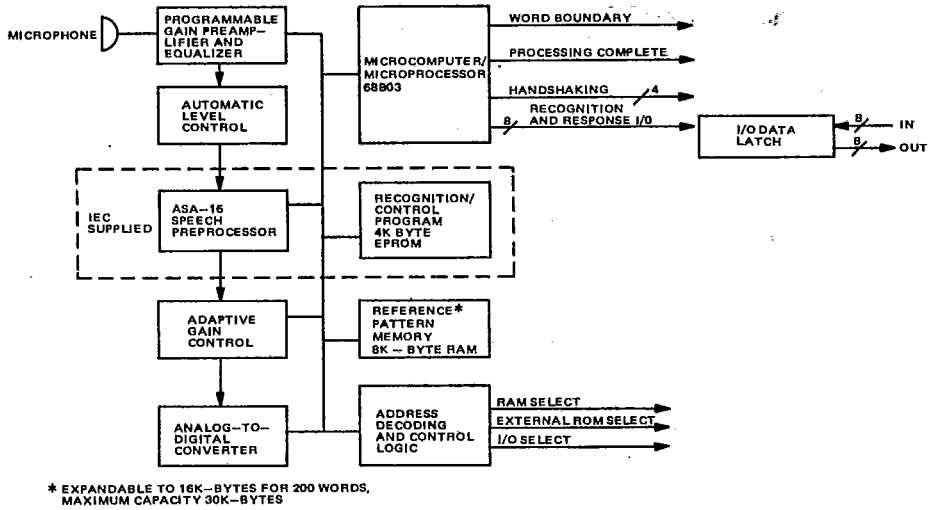


Figure 1. VRC100-2A Simplified Block Diagram

Table 1. Technical Characteristics

Parameter	Description
Power Requirements	
ASA-16 Speech Preprocessor	±10 volts at 40 milliamperes maximum.
Recognition/Control Program	+5 volts at 160 milliamperes maximum.
ASA-16 Audio Characteristics	
Bandwidth	200 to 7000 Hz.
Input Impedance	Greater than 100 kilohms, ac coupled.
Voltage	0 to 7 volts rms maximum input to ASA-16 analog input (at pin 6)
Performance	
Vocabulary Size	Up to 100 isolated words and/or phrases (expandable to 200 words using additional RAM).
Recognition Accuracy	Greater than 98 percent for typical application vocabularies.
Reject Threshold	User programmable.
Utterance Duration	1.25 seconds maximum.
Between-Word Pauses	160 milliseconds (user selectable).
Word Length	Minimum 80 milliseconds (user selectable).
Response Time	(50 + N) milliseconds, where N = active vocabulary size, with a 8-MHz crystal.
Temperature Characteristics	0 to 50°C.

ASA-16 Speech Preprocessor

The ASA-16 chip is a 28-pin integrated circuit with 4,800 equivalent transistors. It provides audio spectrum analysis over the range of intelligibility for speech, 200 to 7000 Hz. (See figure 2 for ASA-16 pin assignments and table 1 for technical characteristics.)

The analog input to the ASA-16 is 7 Vrms maximum, from a low-output impedance source of 600 ohms or less. The ASA-16 consists of 16 bandpass filters each followed by a halfwave rectifier and a second order low-pass filter with 25-Hz cutoff. The monolithic ASA-16 utilizes NMOS switched-capacitor technology with 100 operational amplifiers to achieve the required audio spectrum analysis. Additionally, this chip contains a 16-channel analog multiplexer and decoder and provides all the necessary timing signals from a single TTL 1-MHz clock. Each bandpass filter center frequency is linearly related to the clock frequency. Clock translation results in spectral translation. The analog multiplexer is addressed via four TTL lines. The analog output of the chip is from a buffer amplifier. This output is suitable for input to a 9- to -5 volt user-supplied analog-to-digital converter. (National part number ADC0804).

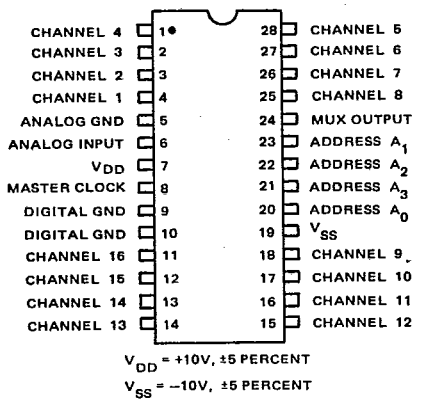


Figure 2. ASA-16 Speech Preprocessor Pin Assignments

VRC100-2A Recognition/Control Program (EPROM)

The recognition/control program is a 24-pin semiconductor device which contains the entire algorithm for recognition of isolated speech utterances including: (1) word boundary detection; (2) spectral slope coding and binary sonogram coding; (3) end-point time compression; (4) template comparison with an input utterance; (5) final decision algorithm; and (6) system control function. The chip provides control of the ASA-16 analog multiplexer and a user provided analog-to-digital converter in conjunction with the Motorola 68B03 microcomputer/microprocessor. The spectral slope coding preserves the frequency versus time characteristics. The binary sonogram preserves the energy versus time characteristics.

The input and output signals for the recognition/control program are illustrated in figure 3 and listed in table 2. Assigned memory map locations for the recognition/control program are as follows:

Internal RAM	0080 through OOFF
External RAM (30K-bytes maximum)	0800 through 7FFF
Multiplexer Address	8000
A/D Start	8010
A/D Read	8020
Set DAC Count	8030
Set Preamp Gain	8040
Set Bias	8050
Read Data Register ($\overline{DOUT}$)	8060
Write Data Register ($\overline{DIN}$)	8070
Optional ACIA 1	9000
Optional ACIA 2	A000
Optional VIA	B000
External ROM/EPROM (16K-bytes maximum)	C000 through FFF (mode 2)

Table 2. ASA-16 Input and Output Signals

Signal	Pin No.	Signal Description
Channel 1 through 16	1 through 4 11 through 18 25 through 28	These pins provide output connections for each of the 16 spectrum analyzer channels prior to multiplexing. The high impedance outputs should be loaded with more than 200 kilohms. These low pass filter outputs permit using an external multiplexer and A/D converter such as the National ADC0816.
Analog Input	6	The speech input is applied to this pin following microphone amplification. The input signal level should not exceed 7 volts rms.
V_{DD} and V_{SS}	7,19	Power is supplied to the ASA-16 using these pins. V_{DD} is +10 volts and V_{SS} is −10 volts, ±5 percent.
Master Clock	8	The master clock is a 1-MHz input signal that synchronizes the ASA-16 logic.
Multiplexer Output	24	The ASA-16 on-board analog multiplexer output is available at this pin. The output voltage is 4.5 volts dc ±10 percent for a 5-volt rms signal at the analog input pin at the center frequency of the corresponding selected multiplexer channel.
Address (A_0 through A_3)	20 through 23	The control signals applied select the multiplexer channel for output from the ASA-16 onboard analog multiplexer. Pin 21 is the MSB and pin 20 is the LSB.
Digital Ground	9	This line should be connected to a low TTL logic level. With this input low, the analog multiplexer output corresponds to the channel specified by the active 4-bit multiplexer address.
Offset Adjust	5	This pin provides a method for compensation of the ASA-16 offset characteristics.

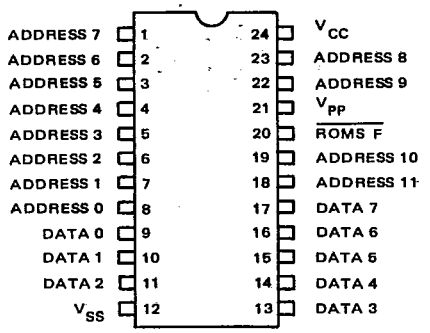
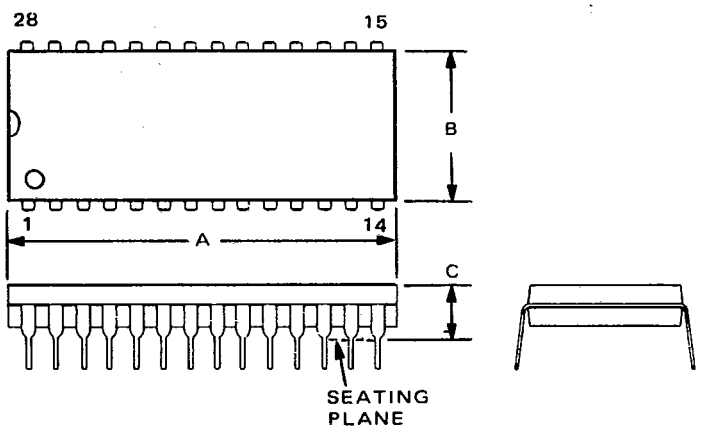


Figure 3. VRC100-2A Recognition/Control Program Pin Assignments

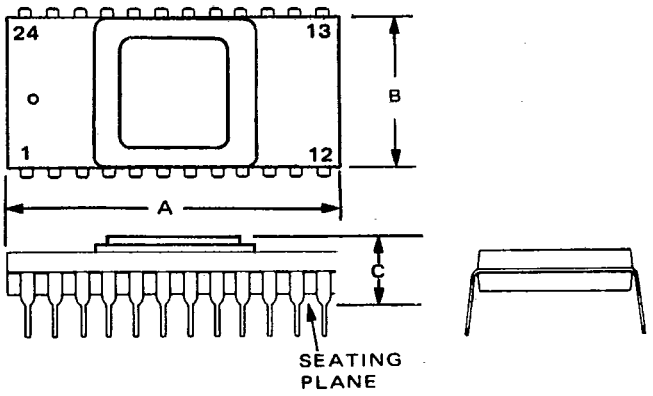
Packaging

The VRC100-2A chip set is available in plastic (PSUFFIX) packages. Figure 4 illustrates the package dimensions for the ASA-16 speech preprocessor and recognition/control program.

ASA–16
SPEECH PREPROCESSOR



RECOGNITION/CONTROL



ASA–16
SPEECH PREPROCESSOR

RECOGNITION/CONTROL
PROGRAM

	MILLIMETERS		INCHES		MILLIMETERS		INCHES	
DIM	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
A	36.45	37.21	1.435	1.465	29.34	30.86	1.155	1.215
B	13.72	14.22	0.540	0.560	12.70	14.22	0.500	0.560
C	4.57	5.08	0.180	0.200	3.05	3.94	0.120	0.155

Figure 4. Package Dimensions

Typical Circuit Configuration

Typical circuit configuration consists of the following: (1) programmable gain preamplifier; (2) speech spectrum equalizer; (3) linear amplification or ALC signal processing prior to the ASA-16 input; and (4) adaptive gain control. If ALC is not required, considerable component savings may be realized. The post-gain following the ASA-16 provides, via the DAC0832 multiplying DAC, adaptive gain control. If this feature is not utilized, a fixed positive gain between the ASA-16 and A/D input of 1.25 is recommended resulting in component savings. Automatic offset compensation occurs: (1) on a power-up; (2) on a reset command; (3) on a train command; (4) on a recognize command; and (5) if a beginning of word has been detected but no end of word has been detected for at least 3.75 seconds. The ALC output is nominally 0.775 volts rms a U11 pin 7 for inputs from -22 dB to +14 dB relative to 0.775 volts rms. The decoding logic can also include any possible requirement for extended RAM, EPROM, or I/O. When bus-to-bus parallel communications are utilized, the AM2952 8-bit bi-directional I/O port (manufactured by Advanced Micro Devices) minimizes interface components. In this circumstance, the 74LS374's may be replaced by the AM2952. Specific decoding designs can be component minimized by the use of programmable array logic integrated circuits.

INTERSTATE VOICE PRODUCTS
A FIGGIE INTERNATIONAL COMPANY

Eastern Region Office

1735 Jefferson Davis Highway, Suite 705, Arlington, Virginia 22202-3483
Telephone 703/892-1400

Interstate Voice Products

1849 W. Sequoia Avenue, Orange, California 92668
Telephone 714/937-9010 TWX 910-591-1197 Telex 47-22046 or 65-5419