

Speech Recognition Methods for Controlling Cable Television

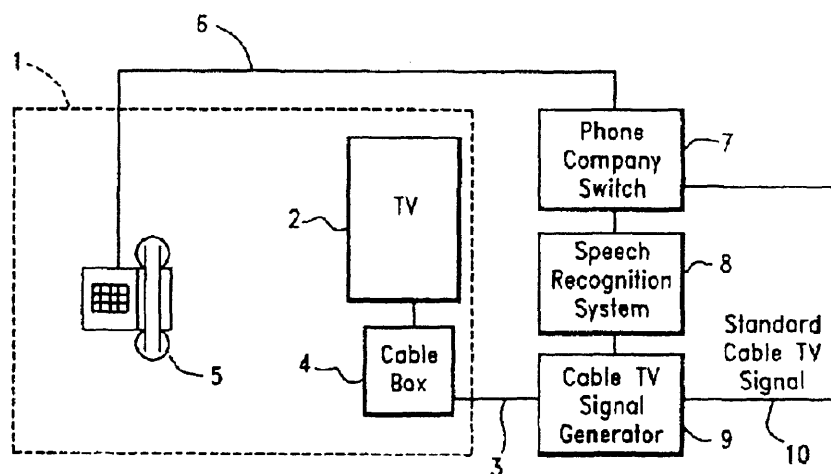


FIG. 1.

Disclosed are methods for using speech recognition to select or modify images, sound, and data transmitted on a cable television system. A first method connects the public telephone network with a speech recognition system and with the cable television network. A second method uses the cable television network to transmit speech commands to a speech recognition system.

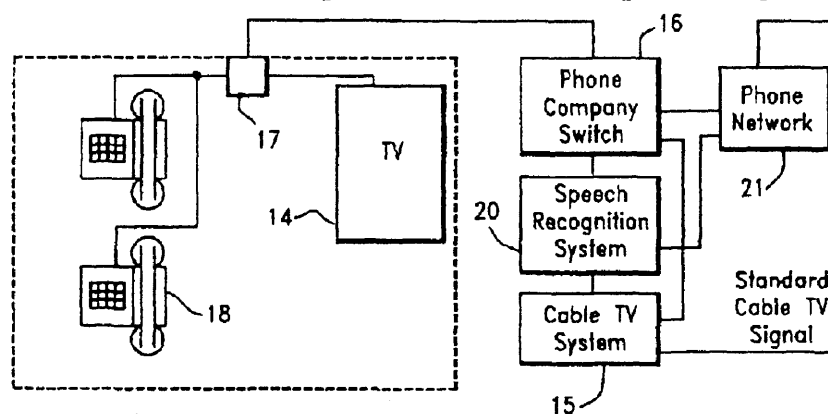


FIG. 2.

Fig. 1 is a diagram of a system in which the first method is applied. A user's house 1 includes a conventional television set 2 connected to a cable television line 3. An optional cable box 4 may be used for this connection. A conventional telephone 5 is connected to a conventional telephone network 6. At a remote location the switching system 7 of the telephone company is connected to a speech recognition system 8, which is in turn connected to a cable television signal generator 9. A standard cable television signal, brought into the signal generator on cable 10, is modified within the signal generator, according to instructions from speech recognition system 8, to form the modified signal transmitted to the user on line 3.

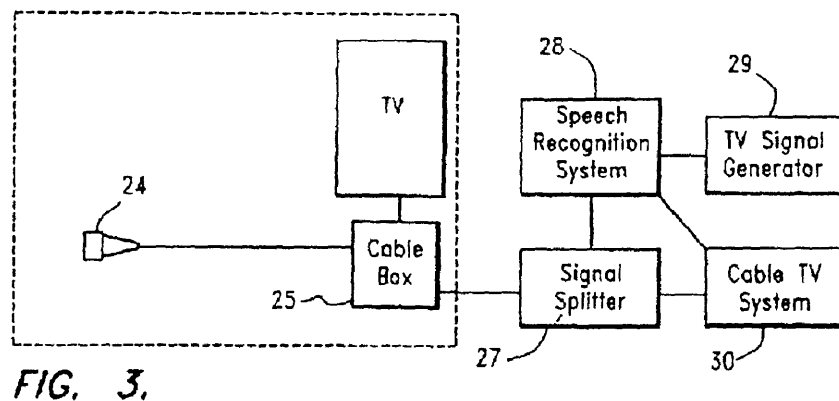


FIG. 3.

The user begins the process of using the system of Fig. 1 by placing a call to the cable television system with telephone 5. The cable system identifies the user with, for example, the Automatic Number Identification (ANI) protocol, or asks the user to provide his phone number or password by speech or by depressing a sequence of telephone buttons. The cable system then displays a series of prompts on the screen of television set 2, on a specific channel or overlaid on all channels. This process may cause the existing channel window to shrink, providing a split-screen mode with the prompts displayed alongside the normal picture received on the channel. The cable system can additionally transmit audio messages to the user through the telephone network. This capability is used, for example, to allow the user to listen to an alternative audio broadcast corresponding to the channel being viewed. Voice commands, interpreted through speech recognition system 8, provide for various functions, such as changing channels, identifying a subset of active channels through which the user may move, viewing multiple channels which are windowed or tiled, displaying text on the screen, selecting answers to multiple-choice questions, seeing scores or personal game totals displayed on the screen, or selecting movies to review.

Within speech recognition system 8, the active words, phrases, n-grams, and grammars are constrained to a minimum number sufficient to support a given screen, panel, or set of prompts. The active context is dynamically changed by user and program input. The resulting images, sounds, and prompts are displayed on the screen means of techniques such as compression, individually addressed television sets, and signaling the cable television to change channels to a public or private channel or to pull selected images, text, or audio from a shared channel. Using such techniques, hundreds of images can be transmitted on a single channel, with a specialized television set or cable control unit 4 holding shared data. For example, the answer to a common question, such as asking what is on television tonight, is broadcast every few seconds, so that the cable control unit can grab the answer as it goes by. Certain audio commands instead require that a specific message be sent to the user's television set 2. Pointers may be used to answer questions destined for many users during game shows, educational programs, and action adventure shows. With this technique, thousands of users can control their own experiences with a small number of channels or sub-channels.

Alternately, the speech recognition system 8 may be used to control a cable television system providing shared video and private audio, in which the video signal portion of a television signal, which may include moving or still images, is common to many subscribers, while the audio portion is unique to an individual user. Since the audio portion consumes significantly less

bandwidth than the video portion, a large number of audio signals may be easily transmitted for targeting to individual users.

If the cable control unit 4 has adequate memory, the screens, objects, images, and sounds associated with information frequently accessed by a particular user may be broadcast infrequently and held for subsequent access by the user. Furthermore, a control unit 4 with adequate memory may be used to record an entire program, providing sophisticated television recording functions that can be accessed remotely.

Fig. 2 is a diagram showing an alternate system in which the first method is applied. A television set 14 is connected to a cable television system 15 through the telephone company switch 16. A signal splitter 17 directs the appropriate signals to television set 14 and to various telephones 18. The operation of this system is similar to that of the system described in reference to Fig. 1, with voice commands being directed through the telephone company switch 19 to a speech recognition system 20, which controls the signal provided from cable television system 15. Conventional connections are also made to other portions 21 of the telephone network.

Fig. 3 is a diagram showing a system in which the second method is applied. Speech recognition is integrated into a cable television system to control the information transmitted to an individual user. A microphone 24 is connected to a specialized cable convertor box 25, providing for control of the cable television system according to voice commands given by the user. From the cable box 25, acoustic data from the user is sent over the cable TV network 26, having been decoded partially or entirely in the cable box 25 or in the cable gateway. Specifically, the user's voice commands are passed unaltered onto a channel or sub-channel of the cable television system, are compressed prior to transmission, are intermixed with transmissions from other users, or are labeled or partly decoded in the home before transmission. In general, voice commands are directed from a signal splitter 27 to a speech recognition system 28, which in turn controls a television signal generator 29 and a portion of cable system 30. Alternately, voice commands are completely decoded in the cable convertor, which then transmits only commands and requests for information on the cable television network.

A wired or wireless telephone handset, connected to the cable box 25, is alternately used in place of microphone 24. A variation of cable box 25 is connected to both the cable television network and to a telephone network, allowing the use of conventional telephones and high-quality digital or spread-spectrum portable phones for both telephone applications and cable television applications.