Voice Of Authority

Speech recognition systems are being increasingly used in commercial and security applications. However, their powers are restricted by the computer's memory capacity

For a computer to be of any use it must have a workable means of allowing commands and information to be fed into it. The 'interface' that we normally use to communicate with a home computer is a keyboard (though mice and joysticks are possible alternatives). By using a keyboard, however, we find that we are forced to communicate with the system by means of an artificial language. Commands such as CLS, DIRECTORY, RUN, LOAD and SAVE may be meaningful to the operating system but they aren't 'natural'.

The natural communication system for humans is speech, not typing messages on keyboards and watching the replies on television sets. If a computer could be made to understand spoken commands — even if they were phrased in the same way as the ones given through a keyboard — it would be far easier to use, especially by those with a physical handicap. Before any computer system can 'understand' spoken words, it must first process the sound input: the analogue signals must be analysed and turned into a digital form that the computer can deal with. Although it seems to be an easy thing to generate electronically, speech is a remarkably complex combination of sounds.

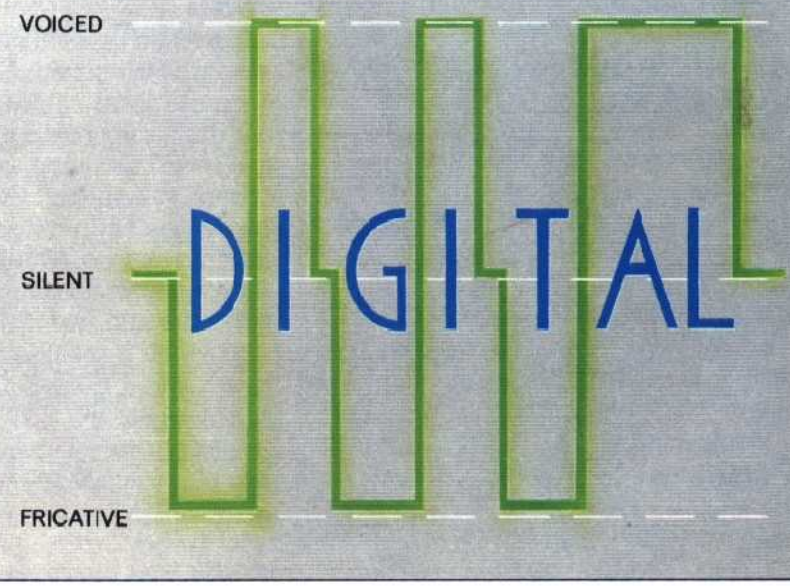
Dreams of instant and complete speech recognition (as typified by the computer HAL in *2001 — A Space Odyssey*) are unlikely to be fulfilled for many years yet, if ever. The voice input typewriter is equally distant; yet the technology for both this and the 'understanding' computer already exists. But neither is available at low cost, because there is a major difficulty in creating speech recognition systems: words can sound the same but have different meanings, depending on the context that they appear in. The processing power needed to solve this problem is simply not available at a reasonable price.

Although researchers have created systems that approach this goal, they have discovered that increasing the number of speakers who can be recognised by the computer has the effect of reducing the number of words that can be recognised at any one time. Typically, a multi-speaker recognition system will allow between 20 and 30 words to be recognised at a time, with a success rate of around 85 to 90 per cent.

The potential uses of speech recognition systems are considerable. The German Post Office uses one to assist with sorting mail; and there are now many applications in aerospace, both military and civil, where pilots have literally not enough hands and feet to control their aeroplanes. In all these situations the number of words that can be recognised at any one time is limited to around 20. However, this doesn't mean that the overall system is restricted. The user is selecting one of the 20 words from a 'menu', and each recognised command produces a further menu of words to choose from. Only when the complete sequence has been successfully recognised will any action be taken by the computer. In the case of the sorting office the first level of sort could be by state, and once the correct state is selected the next sort could

Parts Of Speech

One technique of speech recognition simply involves digitising the signal and performing extensive 'pattern recognition' analysis. A more efficient method is to use hardware pre-processing, in which a number of independent circuits measure the signal for voiced sound (e.g. vowels), fricatives (s,f,t, etc), and short periods of silence (e.g. between syllables). The output from each of these filtering devices is a string of 1s and 0s, which the computer compares with a library of stored examples, selecting the nearest match as the word it recognises



KEVIN JONES