

IMPORTANCE OF SPEECH ON OPTICAL DISKS FOR REAL-TIME APPLICATIONS

P.N.Girija

ABSTRACT

The demand is increasing day by day for systems that can store, retrieve and manipulate such as data, text, graphics, speech and images. This increasing demand has been caused by the rapid development of workstation that has made sophisticated workstations available at low costs. In this article, the stress has been made to focus the efficiency of speech as an I/O interface in intermediate multimedia systems.

Introduction :

Speech as an input/output has been several advantages. Speech is most natural communication mode for human beings and hence much training is not required. It has been proved that speech is faster compared to typing. The key to use speech in Multimedia systems is that it increases effectiveness of communications not only between computers but also between people. Insertion of speech card into computer facilitates in several useful applications.

Speech is remarkable for variety of rules it follows and even more remarkable for the rules it violates. Written sentences may be expected to obey most grammatical rules and still communicate. These are several advances in the PC platform such as voice mail, voice messaging and generating interest in computer telephony integration.

Among current favourites is multimedia which means combining words and numbers with images and sound to create more comprehensive and hopefully more effective communication between computers and their users. Computer support of human interaction with multimedia information is rapidly accelerating, as evidenced by the phenomenal growth in use of digital technology for speech processing and the advances in computer graphics. When fully digital multimedia systems are readily available, we have powerful tools for improving human-human collaboration and human-computer symbiosis [1]. High resolution images, high-fidelity speech, nicely typeset text and high definition video are available.

Brewer [2], Musman [3] and Bruno [4] discussed uses of audio as well as technical issues relating to CD-ROM. Three dimensional speech technology, which can generate physical representation of speech, promises more realistic audio simulations[5,9-13]. Speech is the common currency of office life. So, of all the fashionable multimedia hypes, the fundamental change is speech applications. Speech increases the

number of users and greater demand for communication bandwidth. Speech has rich and social orientation of interaction. There is an evidence that the presence of audio channel makes more of a difference than the presence of a visual channel.

Multimedia systems provide variety of perspectives on the emerging field of interactive digital audio. It is hoped that the article will help encouraging future in this field to meet the challenges and opportunities. Since speech occupies larger memory this should be compressed using some coding techniques. Adaptive delta pulse code modulation is the efficient compression technique which is being used. Digital text, graphics, images, video and audio can be played from a single optical disk. Different applications make different demands on the kinds and level of interactivity.

Speech can be digitized and stored on hard disk in real-time. Once speech and images have been captured, various tasks can be reviewed. Due to large amount of multimedia traffic for audio, visual applications these require high speed net works, to receive data in real-time.

Storage technology details have been explained by E.A.Fox[7].

- In 1982, CD-DA (compact disc - digital audio) was launched, adapting videodisc technology for storing upto 72 minutes of high quality digitally encoded speech.
- In 1984, the first prototype CD-ROM (compact disk read only memory) drives were demonstrated, allowing personal computers to access over 600 MB of prerecorded digital data.
- Also in 1984 writable optical disks, often called WORMs (write once read many), became commercially available use for archiving a gigabyte or more of digital multimedia information.
- In 1986, it was announced the CD-I (compact disk interactive) specifications would be submitted that would handle compact disks with video, images,

graphics, audio, text, data.

- In 1987, digital video interactive have been demonstrated. It could decompress 72 minutes of highly compressed full motion video and FM quality audio in real time using special boards in an IBM PC/AT with attached CD-ROM drive.
- In 1988, erasable optical disks became commercially available, storing hundred of megabytes on stable removable media, so that large volumes of digital multimedia can be economically manipulated and readily distributed.
- Also in 1988, an ISO (International Organization for Standardization) standard specifying the volume and file organization of CD-ROMS was approved, allowing inexpensive worldwide publication and distribution of multimedia data that is directly accessible across a wide range of hardware and operating systems.

The placement of audio data on the optical disk is of primary importance because the audio data is being extracted from the optical disk in "real-time"[8].

Sound Blaster Card or Speech Card :

Several types of speech cards are available in the market which have capabilities of both speech recognition and speech synthesis. Since speech can be more efficiently used for real-time applications, it would be more appropriate for library applications. There are two types of speech systems known as speaker dependent systems and speaker independent systems. In speaker dependent systems, the training of the vocabulary by the speaker is compulsory whereas in speaker independent systems, the training is not necessary. But however these systems are expensive compared to speaker dependent systems. Speech recognition task have been performed by giving speech as an input through microphone. The analog signal has been converted into digital and finally by patterns matching techniques, speech patterns of both trained and input would be tested. The patterns which would match better with one of the trained words, would be recognized. In synthesis the text is converted into speech by using Coarticulation and Prosodic rules. A good thing has time of its own. So speech is the future one!

Conclusions :

As the technology grows more powerful and less expensive, it becomes increasingly accessible to business, government and the average citizen. People can create real-time audio environments which is more useful for all sections of people. Since speech can be processed in real-time, speech interface is more useful in libraries which could be used in several sections like acquisition, circulation, serials

control, production of catalogue cards and inter library loan, etc. By adding speech recognition and speech synthesis the multimedia computer becomes capable of processing, all the same media that humans can, except touch and smell.

All most of all multimedia systems usually have sound blaster cards, users can utilize to increase the efficiency of their work. Speech is more user friendly and it can be used as a shortcut instead of pressing multiple keys simultaneously just by a simple command the functions could be performed.

References :

1. Edward A. Fox
Advances in interactive digital multimedia systems, *COMPUTER*, Vol:24, No: 10, 9-21, Oct.1991.
2. B.Brewer
Using audio in CD-ROM, Vol:2, Optical Publishing, S. Ropiequet, J.Einberger and B.Zollick, eds., Microsoft Press, Redmond, Wash., pp: 169-183,1987.
3. H.Musman
The ISO audio coding standard, Proc. Globe Com 90, IEEE, NewYork, 1990.
4. R.Bruno
Compact Disc-interactive images. *Comm. ACM*, Vol: 34, No: 4, PP:131-185, Apr. 1991.
5. A.Desai Narasimhulu
Multimedia Information Systems : the unfolding of a reality, *COMPUTER*, Vol:24, No:10, PP:6-8, Oct.1991.
6. R. Pye and E. Williams
Teleconferencing : Is video valuable or is audio adequate? Telecommunications Policy, PP 230-241, June 1997.
7. E.A.Fox
The coming revolution in interactive digital video. *Comm. ACM*, Vol:32, No:7, PP: 795-810, July,1989.
8. Clement Yu, Wei Sun , Dina Bitton, Qi Yang, Richard Bruno and John Tullis,
Efficient Placement of audio data on Optical disks for real-time applications, *Comm. ACM*, Vol:32, No:7, July.1989.
9. L.Buddina and E. Young
The Brady Guide to CD-ROM, Prentice-Hall,New York, 1987.
10. L.Fujitani
Laser Optical disk, The coming revolution in on-line storage. *Comm. ACM*, Vol:27, No: 6, PP: 546-554, June. 1984.
11. C.Sherman.
The CD-ROM handbook, MGH,New York,1988.