ELECTRONIC PUBLISHING, VOL . 7(3), 117–146 (SEPTEMBER 1994)

A review of multimedia technology and dissemination systems

LEON HARRISON

Department of Computer Science
University of Nottingham
Nottingham
NG7 2RD
UK

SUMMARY

The connotations of 'publishing' are undergoing rapid change as technology itself changes. Placing marks on paper (by whatever means), and distributing the result, are perhaps the first thoughts that the word evokes but nowadays it encompasses an ever-widening range of preparation, presentation and dissemination methods. Video sources, animation, still images and sound samples are now available as methods of imparting knowledge — and all of these are increasingly reliant on technology-dependent delivery systems. The end-user of information contained in such electronic publications has expectations of the delivery and display mechanisms which have been shaped, in the main, by exposure to the broadcast media, whose centrally funded resources are capable of exploiting high-technology solutions. In trying to emulate similar delivery systems at a personal level, the electronic publisher needs to have a general awareness of what present-day technologies can achieve, together with an appreciation of cost and practical issues. This paper gives a brief survey of these newer technologies as seen from today's perspective.

KEY WORDS Multimedia Electronic publishing Delivery systems Standards

1 INTRODUCTION

If the connotations of 'publishing' are changing then 'electronic publishing', whilst adjectivally limited, is every bit as fast-moving. The new technologies are evolving so rapidly that any profound philosophical conclusions about the true nature of publishing are at risk of being invalidated by new presentation and dissemination facilities, which change faster than philosophers can think. For this reason any review of the technology underlying electronic publishing may quickly become out of date. Nevertheless, this paper attempts to illuminate some of the problems that arise as new technology is absorbed into the widening scope of publishing.

In nearly all societies the consumer is bombarded with information from the broadcast media during every waking hour. Much of that information, used to persuade or inform, is presented using the latest technology and most major publishers are aware that their market share will suffer unless they too can bring new technology into their dissemination processes and published products. The decision to purchase published material, for personal or corporate use, is, in principle, a free one. Consumers of the future, and some of the present day, are unlikely to be satisfied with information that is limited to the printed word.

CCC 0894–3982/94/030117–30 *Received 13 August 1993*
 Revised 16 February 1994

This is not to deny that books have many advantages, but reading generally produces a passive learning environment and the author has a strong influence on the logical flow of information transfer. Skipping a few pages can easily render incomprehensible an author's explanation, if this presumes that all previous pages have been read. New technology can certainly support this book-like formality of acquiring knowledge but it can also supply — usually at extra cost — a novel learning environment under the control of the consumer.

The printed word was the technological breakthrough that allowed the mass transfer of the knowledge from one generation to the next. The important words in the last sentence are 'mass transfer'; the ability of large numbers of people to have access to such knowledge is responsible for the explosive growth of technology that has been characteristic of this century.

When the body of knowledge on a subject is small it is possible for an expert in the field to know all that there is to know. As that body of knowledge increases then the expert loses absolute knowledge and has to acquire a referential knowledge base. Far from being a philosophical stance on knowledge, the idea of a referential knowledge base now has its own phrase in the English language. When queried on a fact which is not directly known we often respond by saying, "I don't know, but I know someone who does". The phrase itself implies a confidence that the information exists but is not immediately available and yet, in the modern world, it is almost always the case that, "time is of the essence". Fast access to information from one's desktop can be provided by printed books only when the range of topics is not too large, or when the desk is situated in a comprehensive reference library.

The problems for an information consumer are those of *speed* and *ease of access* to a wide range of sources. The technological solution to these problems is to store, and disseminate, the information using some electronic format. But to understand and use the technology of tomorrow we need to set out a limited, non-referential, knowledge base for those who use books and printed material today. The following sections attempt to do just that, i.e. to discuss possible dissemination methods and to review some of the information formats that need to be supported.

2 DISSEMINATION

2.1 The telephone system

2.1.1 The existing telephone system

The telephone is now an essential form of communication throughout the developed world. Current estimates [1, 2] suggest that there are some 400–700 million telephones worldwide, and that in the USA alone over 800 million calls are made each day. These figures represent an enormous level of connectivity and, potentially, a huge dissemination network for electronic publishing. However, the technology currently used in the vast majority of telephone systems puts severe limitations on the use of this network as a dissemination medium.

At its inception the telephone was intended for voice communication alone. Experimentation showed that the intelligent speech could be conveyed over a telephone system that allowed frequencies up to 3 kHz to be transmitted and received. This 3 kHz range, or *bandwidth*, although acceptable for speech, imposes a limitation on the ability of standard telephone systems to transmit anything other than bandwidth-limited audio. An additional factor is the ratio of the level of the desired signal to the electrical noise on the telephone line. The human brain is outstanding in its ability to be able to recognize speech buried in a large level of noise. Electronic communication equipment cannot, as yet, approach this human ability.

2.1.2 The future telephone system

Planning for a major change in the world's telephone system began as long ago as 1976 [3] but the financial and technological investment in such a large system means that changes are slow to arrive on the desktop. However, the basic structure of the new standards is already in place and currently supports the existing system. The major change required to bring increased facilities to subscribers is to convince the subscribers themselves of the benefits, and then to sell them the new equipment.

Emulation of the old-style telephone exchanges by modern digital equipment is required only for as long as the local loops (the connection from the exchange to the subscriber) are maintained as analogue connections. If the digital domain boundary encompasses the subscriber's location, then a very significant change can take place: the local loop becomes a general-purpose digital communication channel. Multiple services can be integrated via this common communication channel and an Integrated Services Digital Network (ISDN) can be provided. Although ISDN has long been accepted as an acronym it encompasses many ideas. The worldwide standards that must be agreed upon to make the concept a reality have still not been finalized. The International Telegraph and Telephone Consultative Committee (CCITT) is, amongst others, an organization that is involved in the standardization of ISDN. The CCITT ISDN formal description says that the ISDN network provides

> digital end-to-end connectivity through a limited set of network interfaces providing a wide range of service features evolving from the telephone integrated digital network to meet market needs into the twenty-first century.

This laudable aim opens up the concepts shown in Figure 1, and many more. An exhaustive list of possibilities is inappropriate in this paper, but such information is readily available [4]. The two new services most often spoken of are *viewtex* and *teletex*. Teletex is a form of electronic mail whereas viewtex allows interactive access to remote databases and is used, for example, in France, where telephone directories are being replaced by on-line access. Both options require a subscriber to own suitable equipment in order to take advantage of these services.

2.2 Networks

ISDN will certainly bring network access to a much larger user base, both individual and corporate. Once this user base is established then networks will become an attractive possibility for electronic publishing. To put this into context it is necessary to know something about the technology of networks in order to get a feeling for what is on offer now, and what will be available in the near future.

A well-known, idealized, model of networks is the International Standards

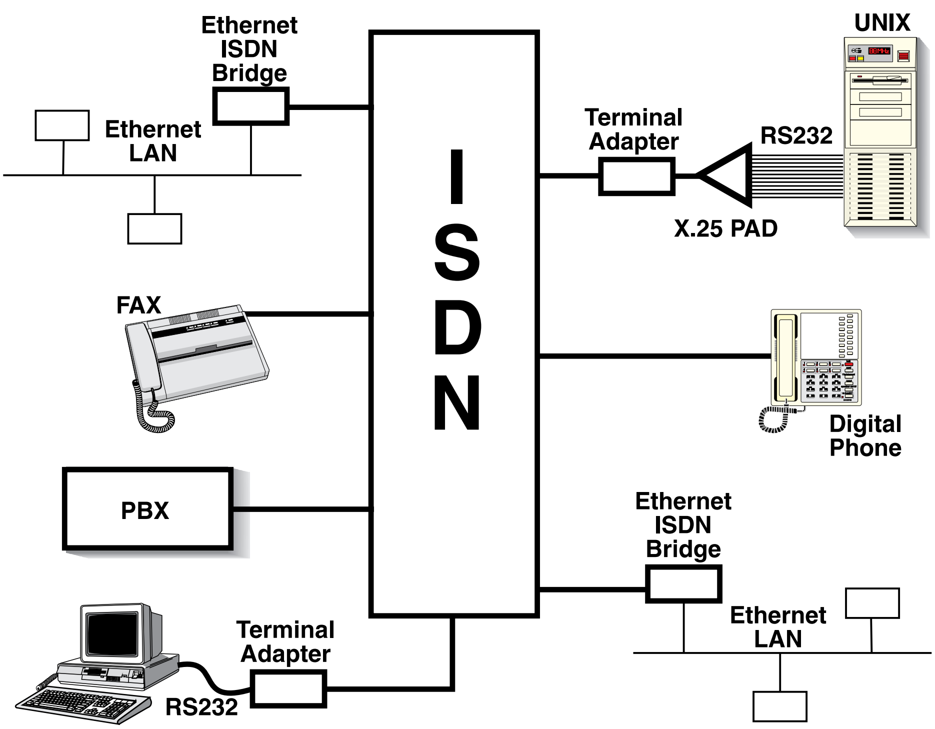

Figure 1. Some ISDN possibilities. Connections to ISDN are at the appropriate data rate

Organization's Open Systems Interconnection layered model (ISO-OSI). In this the network is thought of as a set of seven distinct layers. There are defined interfaces between the layers on a host machine, right down to the lowest layer, which provides the physical connection to another machine, or to an intermediate node. Using this mechanism any layer appears to have a connection with a corresponding layer at another node in the network. The higher the layer the more likely it is that its implementation will be in software. Higher layers are protected from errors in data transmission by the lower layers, and a user of the higher layers, in the absence of significant network delays, may be unaware of working across a network. Unfortunately the physical realization of a network may not conform to the model and the interconnection of non-conforming networks is often difficult.

Some protocols span a number of layers of the ISO-OSI model. X-25, for example, aligns with the lowest three layers and provides an error-free connection for higher layers. It does this by comprehensive error checking as information is passed between intermediate network nodes. However, as network reliability has increased significantly much of this error checking is unnecessary and it slows network performance. Newer protocols, such as Frame Relay and Asynchronous Transfer Mode, ATM, assume that network bit-error rates are small and that the end-points of a connection are capable of sorting out problems that arise from rare errors. Consequently, with higher bandwidth and lower errors, network performance is improving dramatically. The down-side is that more traffic results, and any promised increase in performance can soon be lost.

At the physical level the bandwidth demand is being satisfied by increasing use of fibre optics as a backbone to which copper-based networks are attached. The reliability of such networks can be increased by clever topology which gives built-in protection if parts of the network become inoperative. Ultimately the demands of more powerful workstations will mean that a direct fibre optic connection will be needed.

2.3 The Internet

The Internet is the largest collection of interconnected sites in the world, and it is still growing. Its history can be traced back to the 1970s when an experimental military network, known as ARPAnet, was set up, specifically to do research on network technology and reliability. Because of its military nature one of the concerns was to be able to maintain network operations under catastrophic situations.

In the 1980s the National Science Foundation, NSF, began funding a network using the know-how developed on ARPAnet. The NSF network was firmly in the academic environment and was designed, initially, to allow connection to major computer centres from remote university campuses. Being in the academic area it was not long before the network began to be used for other purposes, such a moving files from site to site, and electronic mail. This NSFnet was being developed at exactly the same time as sites began to use workstations and local area networks. The diversity of hardware used for implementing the local area networks (LANs) created problems when one site, with its own type of LAN, wanted to connect to another site, with a differing LAN technology. Users of a local network wanted all the facilities offered by NSFnet without having to move from their local workstation and it became clear that some standard method was needed for interconnecting networks. The solution adopted on what is now the Internet uses two packet-switching protocols: the Internet Protocol, IP, and the Transport Control Protocol, TCP. These two protocols are nearly always grouped together and referred to as TCP/IP. If a connection is made between two network nodes on widely separated LANs then, provided both use TCP/IP, the connection will allow the two LANs to behave as one. This has allowed the Internet to grow internationally, limited only by the willingness of governments to permit transfer of information across their boundaries, and the provision of international satellite and cable links. Whether the packets cross international boundaries or not they must hop from network to network whilst in transit and the routing of packets needs to be systematic and comprehensive to make sure that all of them arrive at the correct destination. It is the IP protocol which deals with this addressing but it should be noted that an Internet address is the address of a *connection*, and not that of a specific machine or an individual user.

Even if a network does not use TCP/IP it can still join the Internet provided there is a gateway that understands TCP/IP on one side and the local protocols on the other side, and which can translate the information as it passes through the gateway.

Commercial organizations have been part of the Internet for a considerable time, but only to the extent that their research and development departments could be regarded as ‘near-academic’ institutions. Recently there has been pressure to allow direct commercial use of the Internet, with a greater number of organizations being connected, but the exploitation of the Internet in this way is likely to remain limited; there is little likelihood of a transaction on the Internet being interrupted by a commercial !

Figure 2 shows a graphical summary of the nature of the Internet; one of the

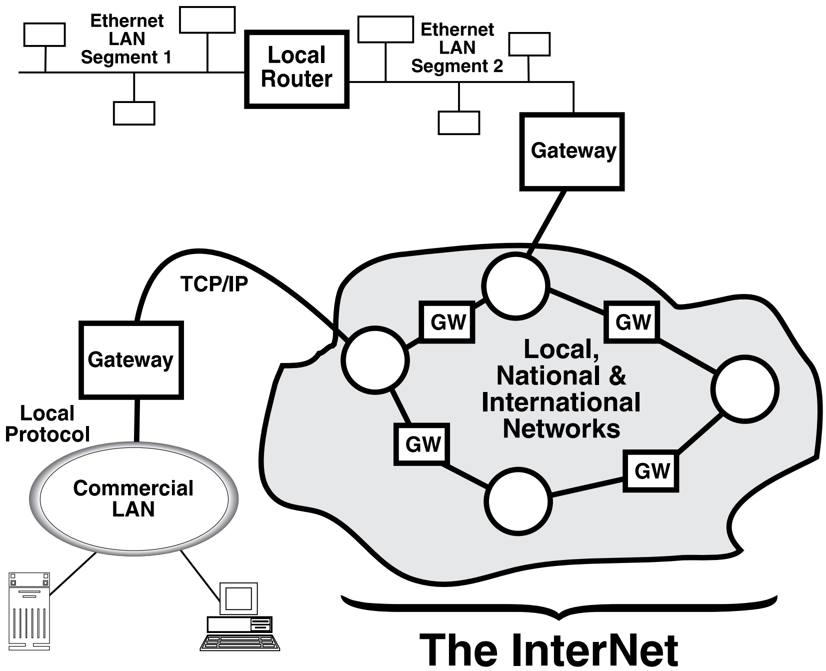

Figure 2. The Internet

surprising things about it is that there is no legal body which *is* the Internet. The organization is essentially run by volunteers, who are usually members of working groups. Joining the Internet is free of charge, except for the cost of a suitable network connection. There are a number of companies around the world which specialize in providing Internet access for individuals or organizations [5].

2.4 Internet services

The importance of the Internet, as a dissemination medium and information resource, is such that it warrants a brief description of the services currently offered.

2.4.1 Telnet

Telnet is the mechanism which allows login access to machines on the Internet but those wishing to use the service must have login rights on the remote machine. Once logged on, the user sees the same interface — graphics or character based — regardless of whether the ‘remote’ machine is in the same room or on the other side of the world. For an electronic publisher it is possible to derive income from copyright material made available on the Internet by simply charging for login rights — provided there is also a willingness to acquire computing machinery, network connections and a system manager to run the service.

2.4.2 Electronic mail

Electronic mail is, perhaps, the Internet's outstanding success. Given that the Internet is just a collection of networks it follows that local electronic mail systems can be made worldwide in scope provided that interoperability problems can be solved. In addition to the sending of personal electronic letters it is also possible to configure an automated *mail server*. If electronic mail is sent to such a server, with an appropriate phrase in the message, an automatic reply can be generated, the content of which is determined by the phrase contained in the incoming electronic mail. This is a possible method for automatically distributing copyright material over the Internet, by validating the incoming electronic mail sender before sending the reply. The mail-server can then initiate the charging method.

Electronic mail is not without its problems, most of which arise because its original design sought to provide an electronic analogue to paper mail. For this reason many mail systems put a limit on the size of message they will transmit based on the plausible assumption that few of us write letters approaching the length of *War and Peace*. Sending messages as multi-part documents can overcome this limitation, but it is far from satisfactory given that the nature of electronic mail will change rapidly as messages increasingly contain images and sound. Under these circumstances, any limitation on message size will have a drastic effect and one must look to protocols such as X.400, which does not limit the size of the message, to remove many of the limitations of current mail systems. In addition, the next generation of mail software will contain some support for multimedia (Multipurpose Internet Mail Extensions — MIME) [6].

2.4.3 FTP

FTP stands for *file transfer protocol* and the name neatly defines its purpose. In common with many other transaction over the Internet, this protocol makes use of servers and clients; an FTP server will usually be a permanently running computer program that responds to request for file transfers originating from an FTP client at a remote location. FTP servers usually require a login name and a password; they can be thought of as a restricted form of dedicated Telnet access. In this form FTP could be used to distribute copyright material — in a similar way to that suggested for Telnet.

2.4.4 Archie

One form of FTP allows access using the login name of *anonymous* and a password consisting of the email address of the client. This *anonymous FTP* is widely used for distributing public-domain information in the academic community, and will presumably be used for advertising purposes as the Internet acquires commercial users. The anonymous FTP sites on the Internet constitute a valuable information resource if some means can be found of indexing them. Archie offers just such a service by organizing a number of sites [7] each of which maintains a database containing the directory listings of all known anonymous FTP sites on the Internet. Logging on to an archie site with a login name of *archie* gives access to software for searching the database using selected key words. By this mechanism information located at anonymous FTP sites can be located quickly and specific FTP sessions can then be initiated to those sites having interesting material.

Clearly, this constitutes a rich source of material for authors and the archie system provides a useful method for distributing public domain information (but it should be noted that copyrighted information cannot be distributed in this way).

2.4.5 Gopher

Gopher is a method of retrieving files from almost anywhere on the Internet, provided that the source of the information is running a Gopher server. The Gopher client, running on a local machine, receives information from the Gopher server about what is available at the server site. The information is arranged, and usually displayed, as a hierarchical file system, with each element of the file system given an attribute indicative of its type. The basic Gopher types are files, directories and search engines of various kinds [8]. If the element of the directory is a file then a simple interaction, initiated by the client, causes the server to send the file. If the entry is a directory then a suitable client interaction causes the content of the directory to be sent to, and ultimately displayed by, the client. It is also possible for search-engine software, at the client end, to initiate interactive sessions with a Gopher server so that specified information can be retrieved.

The original Gopher, named because it could ‘go fer’ information, was designed to be an information provider for the University of Minnesota campus. Rather than one computer needing to contain all the information, the entry in a Gopher directory — which entry is itself a directory — is formatted in such a way that it can be made to point to another Gopher server on a different computer. By making a directory entry contain an Internet address of a remote Gopher server, all of the Internet hosts supporting Gopher can be seen as one large hierarchical directory, limited in size only by the entries in the initial client which the server contacts.

The Gopher protocol is very simple, and does not require the server to retain any knowledge of client interaction. A server simply passes a block of information, a file or a directory, in response to a client request; the client-server contact is then broken. This server’s lack of any *state information* about the transactions under way makes the unadorned Gopher protocol unsuitable for authorization and protection schemes with respect to copyright material. The simplicity of the protocol, on the other hand, is probably responsible for the speed at which Gophers have appeared on the Internet. However, as with all simple things, the simplicity is now seen as a limitation, and the protocol has been expanded to Gopher+, giving enhanced user interaction facilities. The newer protocol is backward-compatible, so old Gophers can still tunnel through the Internet!

2.4.6 WAIS, WWW

WAIS is an acronym for *wide area information server*, which denotes a service for performing indexed searches on documents located on the Internet. The basis from which WAIS works, is a draft ANSI standard, Z39.50, formulated for bibliographic information retrieval. The WAIS server obviously needs access to a suitable index, and then the WAIS client can initiate searches using that index. The result of the search is a list of documents and scores. The higher the score, the better the document matches the search criteria. The search engine employed is not too intelligent as yet, and expertise gained through usage is the best way to get the most out of WAIS.

World-Wide Web (WWW), by contrast, is a hypertext view of the information on the

Internet. The protocol used in WWW is called HTTP (HyperText Transfer Protocol). The price paid for a hypertext view of the Internet lies in the effort needed to prepare hypertext documents, but once this has been done links in the documents can then point to other documents — and so the Web expands. The hypertextual structuring of information and the need for a structured traversal of the Internet means that more effort is needed to to make information available under WWW than would be required, say, in setting up a simple Gopher server.

The concept of WWW started at CERN, the European Particle Physics Laboratory. Information on WWW can be obtained by opening a telnet session to `info.cern.ch` which starts up a client (or *browser* in WWW hypertext jargon). This same IP address can be used as an anonymous FTP source for browser software, both text and graphics based, for a number of computer platforms. WWW documents have the ability to encapsulate material other than text, but the major achievement has been in making traversal of the Internet transparently available to the user when resolving hypertext links. Development is still at an early stage and Gopher+ has some of the attributes that WWW hopes to offer without the need for hypertext documents. Indeed it is possible to make use of Gopher and WAIS servers from within WWW.

2.5 Broadcast and CATV

TV generally is a non-interactive medium, although cable TV, CATV, does have the potential for allowing interaction. In some respects TV output is an archetypal instance of electronic publishing, but some simple calculations show it to be a very inefficient system.

The major colour TV standards, NTSC, PAL and SECAM, including their associated sound channels, need channel spacings of 8 MHz, and, for good picture quality, a signal to noise ratio of 45 dB. Using these numbers, it can be shown that the same channel bandwidth could be used to support a digital link at 120 Mbps. This under-utilization of bandwidth is not the fault of the broadcasting authorities; it is a historical consequence of the way that TV formats were developed.

At the present time the spare capacity of a TV channel can be used by interleaving digital data into time slots in the signal when no screen information is present. Broadcast authorities use one of these time slots to control slave transmitters, and to send test signals. There is still time in this slot, however, to send a relatively slow bit stream containing encoded ASCII data. This is known by the generic name of TELETEXT. It is used on all terrestrial channels, and some direct broadcast satellite channels, in the UK. On commercial channels TELETEXT is used for both advertising and general information distribution. It is not impossible to conceive of publishers using TELETEXT as a method of sending limited forms of material, but modifications to receivers would be required to save a permanent copy.

One commercial channel [9] in the UK has used the broadcast signal directly to send both textual and graphic information. The technique, called *Databurst*, sends, a sequence of still images on a few successive TV frames. The viewer is warned of this, and needs to set up a VCR to record the sequence. At some later time, using the freeze-frame capabilities of the VCR, these still images can be viewed. The limitation of this system, like the TELETEXT option, is the need for some method of providing a more permanent record in the form of a video printer.

One final use of TV as a publishing medium is to encrypt the broadcast signal, and sell decryption equipment to subscribers — a method widely used for satellite broadcasting. Special interest programmes can use the spare capacity of CATV networks to broadcast such material. The high capital cost of setting up terrestrial, satellite and cable distribution networks rules out dedicated systems for publishing or education, unless financed at a governmental level. However, videotape distribution is currently a very cost-effective option.

2.6 CD-ROM

The development of the CD-ROM from its consumer product predecessor (CD digital audio) can be usefully compared to the role of ISDN in bringing the digital domain to the telephone subscriber. CD-ROM brings with it the possibility of using the digital information on a CD in ways that go beyond the mere reproduction of sound. The ability of a CD-ROM to deliver large amounts of data (650 MByte per disc) at a very low price makes it a direct competitor to high bandwidth-network connections. The philosophical difference between the two is discussed in the final section of this paper. Although not discussed in detail here, the basic physical principles of the CD-ROM are similar to those used in videodiscs. Videodiscs, which are usually two-sided — as compared to a single-sided CD — store a digitized version of an analogue video signal; by contrast CDs are single-sided and they store digital data in a variety of formats. Some understanding of CD technology can be gained by studying the layered model shown in Figure 3.

At the lowest level all CDs share common attributes. There is a single spiral track, usually coated in aluminium, which is read, at constant linear velocity, CLV, starting from the centre of the disc. For CLV the read-head position determines the rotational speed of the disc. This, in part, is responsible for the relatively slow access time of CD-ROMs. In the spiral track there is a series of *pits*, and stretches of unaffected aluminium, called *lands*. When scanned with a laser the lands and pits allow digital information to be read back. The information in the pits and lands is encoded in ways that help to compensate for pressing faults and surface blemishes on the finished discs.

The next layer, shown in Figure 3, adds further error detection and correction. The output at this layer is sufficiently error-free to provide the basis for CD-DA, the format for audio discs. A few modifications to this data storage format enables Compact Disc Interactive (CD-I) to provide a variety of digitized sound outputs, which at the highest quality are near to those of CD-DA and compressed video. The ability to mix all of these formats then forms the basis of CD-I multimedia products [10].

The final layer of error detection and correction is the basis of CD-ROM products. With slight modifications CD-XA (Compact Disc eXtended Architecture) can adopt very similar data formats to CD-I, and is, as a result, often referred to as a *bridge* format. However the majority of CD-ROMs conform to the ISO 9660 standard [11] This standard was based on the High Sierra format, whose name came from the hotel, at Lake Tahoe in California, where the initial meetings took place to define the standard. Most platforms can read ISO 9660 format discs, and by making use of certain data areas unused in the base standard, it is possible to produce discs which are platform-friendly without being platform-specific.

Recordable CDs (CD-R) are a recent addition to the range of CD products (they are also known by the abbreviation CD-WO, where the WO stands for Write Once). They

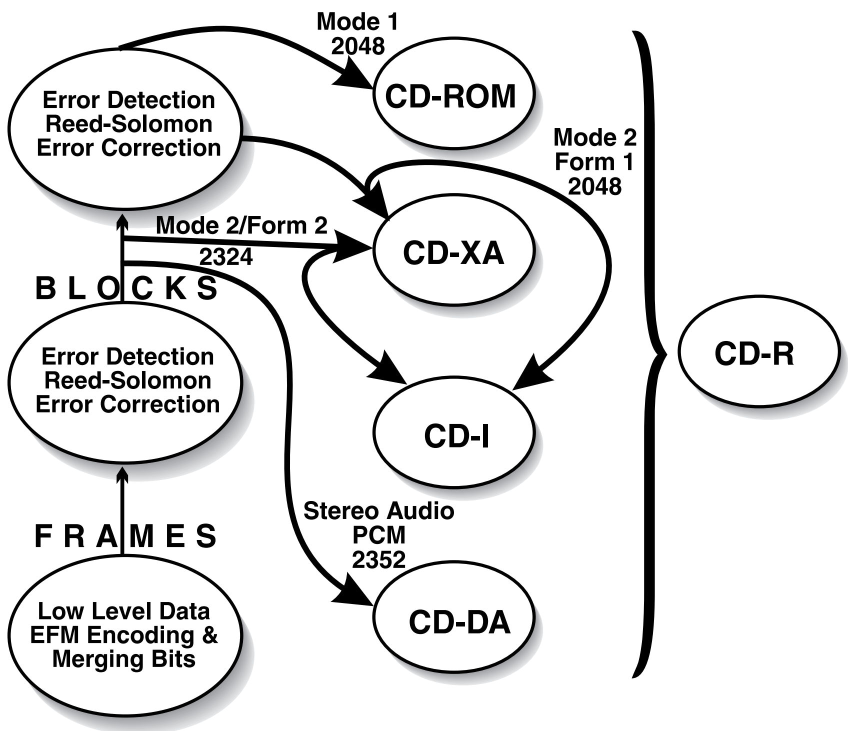

Figure 3. Compact disc formats. Numbers refer to the bytes per sector

are the recording medium used by Kodak's Photo-CD product and this, in turn, has led to blank CD-Rs, and the machinery to record information on them, being available at affordable prices. CD-R allows multi-session recording to take place, and this feature can be used to write more than one *volume* of information on a single CD. With the appropriate hardware and software it is possible to produce any of the types of CDs that have been described, starting from a blank CD-R disc.

2.7 CPUs, memory, disks etc.

The technology described in the previous sections can only be exploited if the cost of the final delivery system is close to consumer-product levels. Electronics is an industry which has successfully created a series of products showing a monotonic trend of cost–benefit advantages. Although research and development costs are high, the economies of scale have, so far, led to low-cost products, even if the benefit side of the equation may sometimes be doubtful. Nowhere is this better illustrated than in the creation of central processing unit (CPU) chips for desktop computers. The mass production of powerful CPU chips has moved from the production of 8-bit models, working at clock speeds of 1 MHz, to 64-bit CPUs, working at 100 MHz, over a time-span of less than twenty years. As a crude measure of performance the product of bit size and clock speed demonstrates how spectacular progress has been.

Interestingly, the rate of progress in CPU development has brought to light a problem

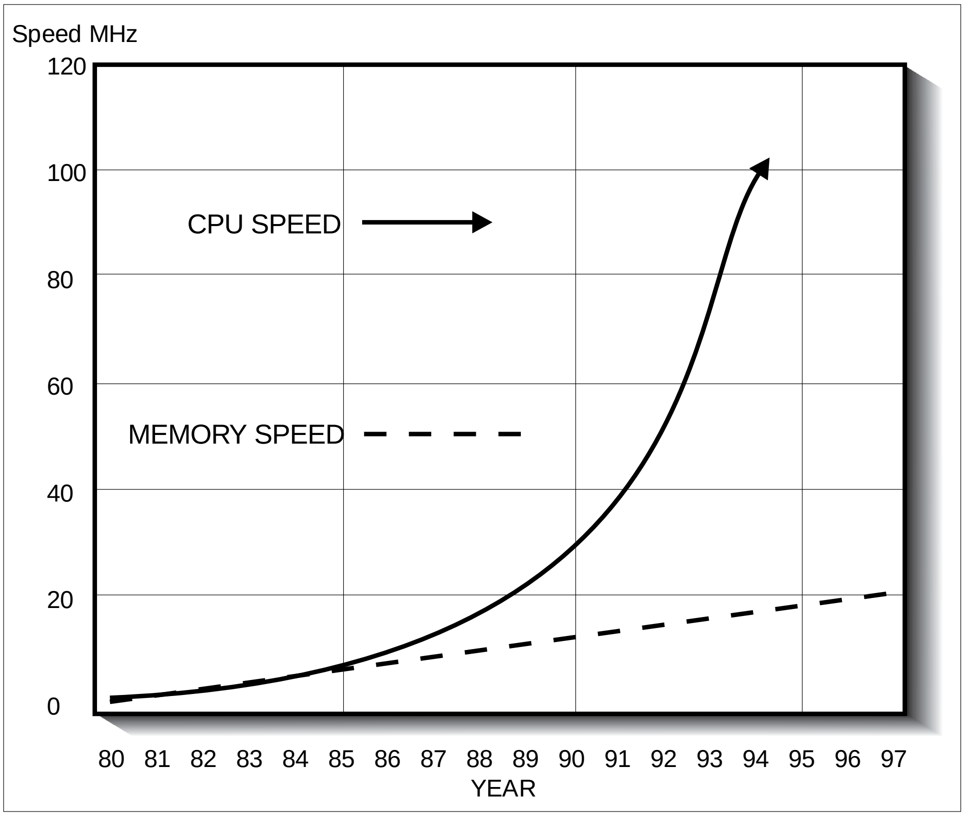

Figure 4. CPU and memory speed comparison

associated with the parallel, but separate, development of memory technology. While the pursuit of higher speed has been the holy grail for CPU designers, memory designers have pursued increasing storage density as the goal for inexpensive main memory. Both design goals have achieved remarkable results, but their divergent aims have caused some problems. This point is illustrated in Figure 4, which shows, on the same scale, a crude plot of CPU clock speed and memory speed (measured in the form of synchronous access speed). As can be seen, CPUs were well served by fast memory systems until the mid-1980s, but nowadays there is a major speed disparity. A part-solution to this disparity lies in the system-level design stage of the computer, and — spreading from there — into the packages used to contain the bulk system memory [12].

Within the choices of system specification for present-day personal computers most of the requirements for fast access to electronic documents can be met. However, the more technologically demanding requirements of multimedia will generally require a greater outlay in the form of CPU power and high-performance peripherals.

2.8 Graphics and displays

Computer graphics is, perhaps, the one area where the greatest technological demands are made on the underlying computer system. On-screen graphical information can be thought of as a representation of that area of memory which contains the information needed to control the intensity and colour of the display. There are, of course, different ways of encoding this information. In a black-and-white display each picture element (pixel) can be represented by 1 bit, the two possible states of which determine whether the pixel is on or off. In a grey-scale image with 256 possible shades of grey, 8 bits per pixel are needed. In a full-colour display the three primary colours need 8 bits each, and hence there are 24 bits per pixel, giving rise to over 16 million possible colours. An intermediate option is to use 8 bits to represent the most appropriate 256 colours for a given application, but with these 256 shades being chosen from the full 24-bit palette of up to 16 million colours.

Whatever colour model we choose we have come to expect high-quality screens on our computer systems. This has only been made possible by adapting mass-market colour television components for use in computer systems. Unfortunately, TV screen formats have been dictated by the TV broadcast standards which, at the moment, all use an aspect ratio of 4 units horizontally by 3 units vertically. The emerging High Definition Standard TV, HDTV, systems will use a 16:9 aspect ratio, but CRTs made to these relative dimensions are still costly. Cost-competitive display terminals will need to stick to the 4:3 ratio for some time yet.

A good definition to work with for detailed images and reasonable quality text is 1024 pixels horizontally, arranged in 768 rows vertically. A 24-bit colour display at this resolution needs 2.25 Mbytes of display memory, while 8-bit colour, or grey-scale, displays need 0.75 Mbytes. In a complete screen update all of this memory has be changed as rapidly as possible and this requires the tightest possible coupling between the CPU and the display memory. In other words, the display memory needs to be part of the CPU's directly addressable memory space, rather than being accessed from a relatively slow back-plane. An even better option is to make use of a dedicated graphics co-processor, to which the main CPU can direct high-level commands. The co-processor itself may, of course, use parallel hardware to implement the high-level commands, thereby achieving very high performance.

In broadcast TV the frame rate is either 25 or 30 frames per second (fps), depending on the TV standard, and each frame is made up of two sequentially transmitted fields containing half of the lines for the complete frame. In one field the odd-numbered lines are transmitted, and in the other field the even-numbered lines. When displaying the complete frame these odd and even lines are interlaced on the screen. For most broadcast scenes this system is appropriate, but for computer displays, where a single, distinct horizontal line may be drawn, the line would appear to flicker at the frame rate if interlaced scanning were used. This leads to considerable viewing fatigue. The solution is to use non-interlaced scanning and a frame rate that is higher than broadcast standards.

The significant differences in bandwidth, frame rates and line display rates, between broadcast TV components and computer displays, is the reason that there is such a price differential between a 21-inch colour TV and a 21-inch colour computer display. To produce a thin-film transistor colour display to match the performance of a similar-sized CRT would require over 6 million perfect transistors to be fabricated. Consequently the CRT will be around for some time to come. However, as standards emerge for HDTV there should be a reduction in prices for computer displays, since many of the requirements for HDTV are similar to those for computer displays. It is unfortunate that the 16:9 aspect ratio for HDTV is no better a match to standard paper sizes than is the current 4:3

ratio. Although the varying shapes and sizes of books on a library shelf demonstrate that there is no such thing as a ‘standard’ paper size, it is still the case that aspect ratios of books are, predominantly, bigger in the vertical dimension than the horizontal — in contrast to television where exactly the opposite is true.

3 PRESENTATION FORMATS

Having reviewed some of the technology that supports the dissemination of information, we now examine some of the encapsulating formats for the representation of information. There are too many such formats for this to be a truly comprehensive survey. However, many of the formats make use of similar techniques so it is hoped that most of the important issues will be covered.

3.1 Text

3.1.1 Character shapes and fonts

Until quite recently any discussion of text presentation. on screen or on paper, would have been very much concerned with font technology. Like many other technologies the ultimate arbiter is the market place and two vector-based font formats — Adobe Type 1 and Truetype — have become well established for use on a wide range of printing devices. The situation with screen fonts is a little more complex, although it is now increasingly the case that modern graphical user interfaces (GUIs) use vector (i.e. outline) fonts rather than fixed-size bitmaps. Many operating systems will now support vector fonts as an integral part of the system’s interface to the windowing software and if the Display PostScript system is used then the same Type 1 typefaces can be used for screen display as well as for high-quality printout. In short, fonts are now an ‘off-the-shelf’ item and those users wishing to design their own character shapes can obtain font design software at modest cost.

3.1.2 Text preparation formats

This area is the publishing equivalent of the Tower of Babel. The plethora of formats has itself produced a mini-industry of conversion software. Faced with the preparation of text, most authors prefer to stay within the computing environment with which they are most familiar, and to rely on the conversion software to translate to the required format at a later stage. Needless to say the conversion is achieved with varying degrees of success, depending on the source and destination formats sharing similar capabilities.

The great divide in text preparation is between formats that describe the ultimate rendering of the text, in clearly visible ways in open text, and those which do the same with hidden codes. The first method needs only the simplest of editors at the preparation stage, whilst the second requires an editor which can generate the hidden codes with appropriate combination of input keystrokes. The second class of editors is now generically referred to as *word processors*. Within this latter class the divide is between What You See Is What You Get, WYSIWYG, and non-WYSIWYG user interfaces. The WYSIWYG concept lies naturally within a GUI, but the GUI offers an extended form of preparation classed as Desk Top Publishing, DTP, with yet another set of file formats, and usually with the software including built-in conversion filters to import from less-powerful word processors. In moving files, other than those of the simplest of text formatting types, the transfer mechanism must be chosen to perform an absolute image transfer. If this is not the case then the recipient of the file may have lost some, if not all of the hidden codes, with the obvious results when opened for further editing.

3.1.3 Page description and document interchange

Over the past ten years Adobe Systems Inc. has succeeded in making PostScript the dominant page-description language. The interpretation of the language is processor-dependent, but the result of the interpretation is device-independent. The majority of problems with the use of the language arise from its specific implementation in a particular hardware environment, and not from the language itself. The language specification [13] and the associated font technology [14] — but not the font designs — are now in the public domain. PostScript is concerned, primarily, with final-form presentation of information on screen or on a printing device; it allows many powerful manipulation techniques to be used in forming this final output. In some respects the power of the language can be abused, leading to unnecessarily complex PostScript output. Most non-trivial PostScript programs will have a prologue, containing local procedures, whose effective use can greatly reduce program size. This is achieved by ensuring that frequently used PostScript-language sequences are set up as procedures with short names. If such a procedure is called in several thousand different places in a program the short-name effect alone will greatly reduce program size.

There are suggested standards for the structure of PostScript programs but the prologue's internal structure is left to the program author's own style of coding. As a result, even an experienced PostScript programmer may find it difficult to understand the body of a PostScript program if it is preceded by a complex prologue. Applications which generate PostScript output use an application-specific prologue, which can be attached to every job sent to a printer. To save downloading the prologue on every occasion, it is possible to send it only once, and to have the printer store the prologue internally. This option is viable only so long as a single application has dedicated use of the output printer. In a shared-printer environment it would be all too easy to find a body of downloaded code interacting adversely with an alien prologue.

In June 1993 Adobe Systems Inc. launched a series of Acrobat products based on a new Portable Document Format, PDF [15]. A PDF file is an electronic representation of a document whose content, and imaging model are derived from Level 2 PostScript. Despite this, a PDF file cannot be sent directly to a PostScript printer since, in some respects, it can be regarded as a highly structured PostScript program body but with no prologue. The missing prologue is similar, but not identical, to that used in the Adobe Illustrator™ package; it uses a great many of the short-procedure-name tricks — elaborated earlier — for keeping file size under control. As a result, unnecessary program complexity is avoided.

PDF allows additional functionality beyond that of simply describing the page. Hypertext links and other annotations are possible in release 1 of PDF; these will take on increasing importance in future releases as they are enhanced to include other objects, such as sound and video clips. Clearly, when material of this sort is present, it needs more than just a simple prologue in order to convert a PDF file into a PostScript program

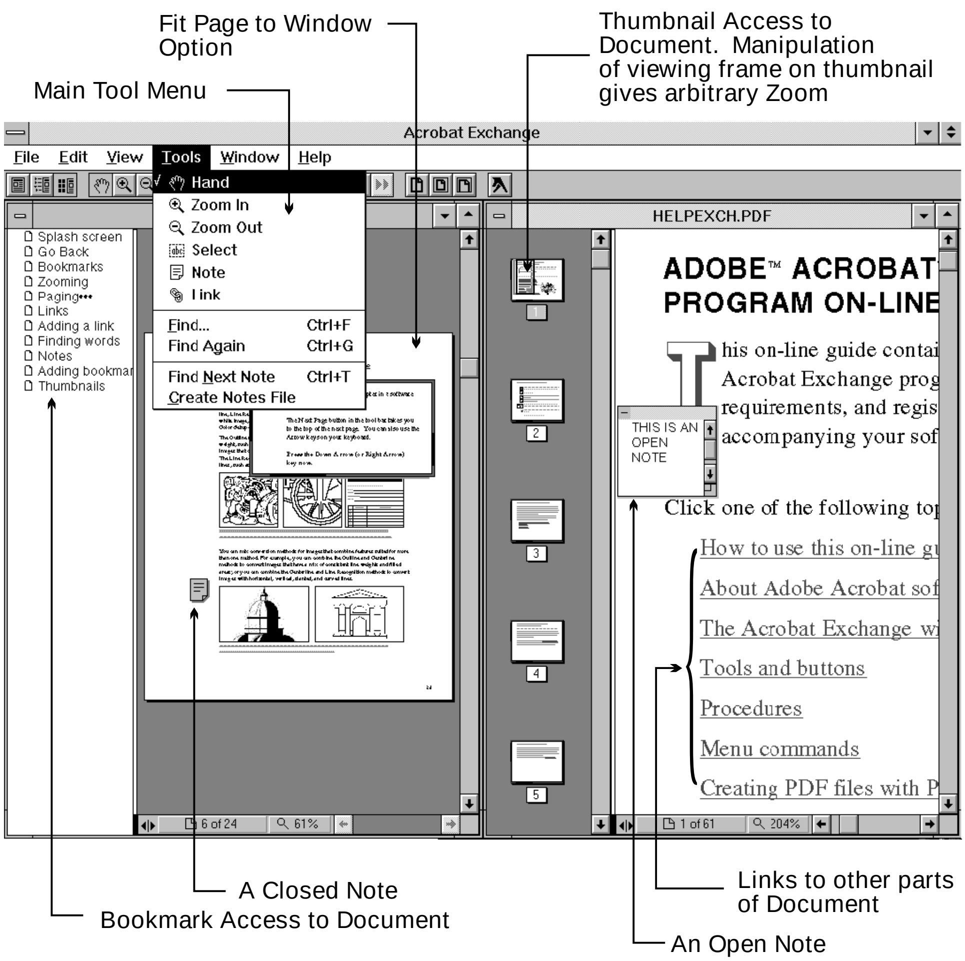

Figure 5. Acrobat Exchange main features

but an even bigger problem is that multimedia inserts will make it hard to maintain PDF documents that are independent of any particular hardware platform. At present, Acrobat maintains platform independence by providing platform-specific software for file-creation, viewing, printing, annotating and hypertext linking — all based on an identical PDF file format. The *atomic* level of the PDF viewers is the page, and the software packages so far released allow the insertion of pages into an existing document, the addition of annotation marks, links to other sections of the document, but not the direct editing of the document itself. Figure 5 shows two separate documents being viewed, with the majority of features of the viewer displayed.

A document exchange format relies for its success on the ease of producing material in that format. Adobe has addressed this problem by providing two methods of creating PDF documents, both of which are available in the Mac and PC GUI environments. The first piece of software is the *PDF writer*, which functions as an extra ‘printer driver’ and is, logically speaking, just another output device. As such, it can be used from any Macintosh or PC software that uses the standard printer interface so as to create a PDF file containing the application's text and graphic objects. This PDF file can then be viewed in an Acrobat viewer. The second software product is the *Distiller*, which, as its name implies, takes input in the form of PostScript code, generated by an application, or by hand, and distils the essence of the program — the page marking — into a PDF file. This is a far from simple task, because of the richness of the PostScript language. A mark on the paper can have been made in an infinite number of ways and deciding what mark a sequence of PostScript commands will produce means that the Distiller must contain the essence of a Level 2 PostScript interpreter, together with the ability to reverse-engineer the page marking into the simplest PostScript coding that is equally capable of producing the given effect.

While the simple description of PDF, so far given, is sufficient for an overview, there are one or two important problems that now require us to go into a little more detail. For example, if a PDF document requires unusual fonts these may not be present on the viewing platform and the simple-minded substitution of fixed-pitch font such as Courier would not give a true representation of the original document. A structure within the PDF file contains information on font name, character widths and style. If the font is present in the viewer environment, then this information is not used. However, if the font is *not* present, then the internal structure is used, in conjunction with Multiple Master fonts, to create a view of the document that is a close approximation to the original. This approximation matches line breaks and pagination but, of course, at high magnifications, detailed knowledge of the original font will allow character differences to be seen. If absolute rendering is important it is possible to embed a type 1 font within a PDF file.

Viewing a document on-screen imposes different requirements from those needed when printing the same document. Serial accesss to the file is sufficient for printing whereas viewing requires random access. Moving from one page to another in a document should not be dependent on searching for page boundaries, as would be the case in serial access. In response to this requirement a cross-reference table in the PDF file provides a random access look up facility, which gives equal page access time, even when moving from first to last page. One final facility worthy of mention in this brief overview of PDF, is compression. The verbosity of PostScript can lead to large file sizes and PDF goes some way towards easing this problem by providing a number of optional compression regimes. Image data can make use of Group 4 FAX compression (lossless) and JPEG compression (lossy); these techniques are covered in the 'still images' section later on in this paper. In addition, the lossless Lempel-Ziv-Welch (LZW) compression method can be applied to the whole of a PDF file.

3.1.4 Printing

Electronic publishing ought to be able to effect a reduction in the usage of paper but if one compares the ease of use of an electronic display against that of paper then the equation is still heavily biased in favour of paper. Most people, when given a choice of reading a lengthy document on a screen, or on paper, will opt for the latter. What electronic publishing does provide, for the printed word, is the ability to generate sophisticated, and complex page styles, which in turn place heavy demands on the mass printer market. Commercial printing of colour and grey-scale images uses dot densities of at least 1200 dpi, and the associated leading-edge technologies, while of interest, affect the end-user mainly by providing an end-product of ever-increasing quality. This leading-edge technology tends to filter down rather slowly to the mass market as an influence on printer design; more significantly, high-quality commercial output raises the desire of end users for creating a close approximation to that quality, on the desktop, particularly in the area of colour printing.

The laser printer in all its forms has been the mainstay of quality, non-commercial printing and has offered a baseline capability of 300 dpi for a considerable time. We can view the laser printer, quite legitimately, as a special-purpose computing machine, with dedicated I/O, and all that has been said on CPU, memory and disc capabilities applies with equal force to such printers. Consequently, performance has improved significantly in terms of a device's ability to translate the native printer control language, speedily, into visible results. Of particular note is that the decreasing cost of memory, and the increase in dedicated CPU speed, has established a new baseline performance of 600 dpi, whilst maintaining throughput. Anti-aliasing, and other enhancement techniques, can now give a performance similar to 1200 dpi. The barrier to even finer resolution may lie in the problems associated with the fine size of the toner particles and the use of laser printers in a normal office environment, although worries on ozone emission from the corona wire seem to have been exaggerated, except in confined spaces. Colour laser printers, to date, have not had a large market share, owing to the high cost of initial purchase and competition from other, less expensive, printing technologies.

Inkjet technology, which offers resolutions around 300 dpi, is the most affordable colour printing technology for the desk top, its biggest disadvantage being the inherent use of wet inks. Some wax-transfer printers, using molten wax instead of ink, overcome the wetness problem at the expense of requiring a heated reservoir for the molten wax. A simpler, and less expensive, wax-transfer mechanism is to make use of ribbons which pass over a thermally activated print-head. Wax-transfer methods can print on many surfaces and special paper is required only when optimum results are absolutely essential. Colour rendering has to make use of halftones in both of the types of colour printers that we have just discussed. For this reason the spatial colour resolution is reduced well below the dot resolution. Dye sublimation printers get around this problem by creating the desired colour on a pixel-by-pixel basis; they offer the best spatial colour resolution and can produce results of photographic quality. However, none of these printing technologies is ideal and, for individuals, the cost can be high ($4 for an A4-size dye-sublimation print) with mass-production of copies being problematic. Nevertheless colour foils for overhead projectors bring an extra dimension of visual impact that is well worth the price.

3.2 Sound

The use of sound in electronic publishing, like animation and video, is problematic in that there is a a shortage of platform-independent formats. To add to the confusion many of the formats are dependent not only on the particular workstation but also on the purchase of specific 'add-on' cards.

Sound formats fall into one of two broad categories. Firstly, there are those formats which do not contain a direct representation of the sound, but instead consist of encoded commands to specific hardware. Secondly there are those formats which contain sampled, digitized sound in some encapsulation of pulse code modulation (PCM).

3.2.1 Non-PCM sound

In the first category the specific hardware determines the nature of the format, and, as a result of the growth in the number of chips now being produced, it is impossible to attempt a comprehensive survey within the space constraints of this paper. One format is worthy of a mention, since it does offer some cross-platform capability. The format is based on the Musical Instruments Digital Interface, MIDI. Since the MIDI system owes no allegiance to any operating system, then any coding of the signal used for MIDI can, in theory, be converted for use on any platform. The physical implementation of a MIDI system uses a serial interface to link musical instruments. The MIDI signals determine which musical note should be sounded and, to a varying degree, the timbre, attack, decay and other qualitative aspects of the note. Many of the qualitative aspects of the MIDI signals are instrument-specific, but the note value itself should be understood by any MIDI device. Therein lies its usefulness. By incorporating the chips used in electronic synthesizers on a suitable card, the card can be interfaced, via a suitable driver, and sent the contents of a MIDI file. The result is that the host computer has the ability to play electronic music. If a full MIDI interface is incorporated on the same card then the computer can be used to control external MIDI devices, and also to record the output of those external devices. This is the basic requirement for a MIDI sequencer, which, with suitable software, can be used to compose or edit musical scores and thus to effect the electronic publishing of music.

3.2.2 PCM sound

PCM-encoded sound is subject to all the limitations of any sampled, continuous signal. In simple terms, for adequate reproduction, the sampling frequency must be double that of the highest frequency to be sampled. The sampling of signals is performed by analogue-to-digital converters, ADCs, and the reconstruction is performed by digital-to-analogue converters, DACs. The number of bits used in the ADC is a direct measure of the resolution of the process. Increasing the resolution of the ADC, by using more bits, will reduce that form of noise which is produced by the *quantizing* of the smooth input slope. The name for this effect is *quantizing noise*. For these reasons, when judging PCM sound, the two major factors to consider are the sampling rate, and the number of bits used during the sampling. If compression is used on the stored PCM file, then the performance of the machine doing the decompression must be of adequate performance to ensure real-time playback.

The results of this discussion are not just limited to sound; they are generally applicable to all systems that use combinations of ADCs and DACs with some intermediate digital storage. For example, images sampled from live TV, after appropriate manipulation, are presented to an ADC/DAC system in exactly the same way as a sound signal. The quantization noise is then perceived as graininess in the reconstructed picture, and the sampling rate places a limit on the fine detail that can be seen.

3.3 Still images

It has been said that a picture is worth a thousand words but, in computer terms, the quality of a picture would not be very good if it required only the same storage space as a thousand words of text! Many image-file formats attempt to resolve this problem by using compression techniques, while a minority sacrifice space in order to maximize retrieval speed. The diversity of image-file formats means that only a limited discussion is possible here. However, an archie search [7] on the Internet will usually provide detailed information on any file format that is in the public domain. We limit our discussion, in this paper, to those compression methods that are used in a large number of file formats. The two major categories of these compression methods are *lossy* and *lossless*. Lossless compression maintains data integrity through the compression/decompression process. LZW, for example, can be used for lossless compression, and, depending on the image content, can produce good compression ratios. Lossy compression, on the other hand, means that data is not preserved, and a decompressed image differs from the original. The human visual system is the arbiter in judging the consequence of the incurred losses, and as a result the compression methods attempt to preserve the detail which is important for correct perception of the image content.

3.3.1 Fax

Fax formats [16] use lossless compression and, apart from their obvious use in telecommunications, they are also used in some Document Image Processing (DIP) packages. Group 3 and 4 fax use the same basic compression methods, but differ in their intended use. Group 3, the predominant format at present, was specified to enable transmission of information over existing, analogue telephone systems, whereas group 4, which offers higher scanning resolutions up to 400 dpi, is aimed at digital ISDN transmission. Grey levels are not directly supported. Group 4 fax is also part of the Document Architecture Transfer and Manipulation, DATM, recommendations from CCITT and ISO. The specification and ratification of such standards is so slow that it is very likely that commercial initiatives, such as PDF, which offer better-quality documents, will ultimately be more significant in document interchange.

Compression in fax is achieved by using Modified Huffman run-length encoding, MH. A scan line across the page is a sequence, or *run*, of either black or white pixels. The length of the run varies depending on the particular scan line, and the content of the page as a whole. Rather than using the runs directly, each black or white run length is given a unique code, and it is that code which is stored. To increase compression the most frequently occurring run is encoded using the minimum number of bits, with less-frequent run lengths being assigned a greater number of bits, up to a maximum of 9 bits. To accommodate long run lengths, other codes are also assigned for those lengths whose sizes are integer multiples of 64. Optimizing the compression would need the whole page to be scanned and stored, and a statistical analysis of run lengths to be performed. The derived encoding table would then have to be stored with the encoded page. However, CCITT has produced standard encoding tables, derived from scanning a selection of pages of text in different languages, and one of these tables is used, rather than a bespoke encoding table for each page. At the end of each scan line an end-of-line code is used, and the bit which follows the end-of-line code signals how the next scan line is encoded. If the bit is 1 then the encoding just described is used. If the bit is 0 then an encoding scheme is employed which makes use of the vertical correlation down a page.

3.3.2 JPEG

JPEG is an acronym standing for Joint Photographic Experts Group. The 'Joint' part of the name refers to the fact that the group was within the framework of both the CCITT and ISO, and thus JPEG standards [17] are common to both organizations. JPEG compression applies to continuous-tone grey-scale, or colour images; it cannot be used directly on bi-level images. The JPEG specification encompasses an interchange format, together with lossless and lossy compression algorithms, but only the baseline lossy algorithm will be described here, as it is by far the most commonly encountered. The interchange format, as described in the standard, is insufficient to specify a complete file format for JPEG-compressed images, particularly colour. As a result, a grouping of hardware and software companies met in March 1991 and established the JPEG File Interchange Format, JFIF.

Baseline encoding uses the Forward Discrete Cosine Transform, FDCT, and decoding uses the Inverse Discrete Cosine Transform, IDCT, which, from now on, will both be referred to as DCT, with the context indicating which of the two is appropriate. The use of the DCT was determined solely by subjective viewing tests, comparing results obtained from a number of transforms. In the absence of rounding errors, the DCT is completely reversible, so it is not the use of the DCT itself that produces the compression in JPEG. The DCT is one of a number of related transforms, which includes the Fourier transform. For the purpose of this discussion it can be thought of as transforming the sampled image from the spatial domain, into the frequency domain. In the JPEG algorithm the two-dimensional DCT is used on an 8 pixel by 8 pixel block of the image. The whole image is scanned horizontally in rows, from top right to bottom left, in blocks of this size, with the DCT being applied to each block in turn.

The result of applying the DCT to the 64 pixels in the image block, is a further block of 64 DCT coefficients. As explained in the section dealing with PCM sound, the frequency distribution is a representation of the detail in an image and this is important for the discussion that now follows. The two-dimensional DCT is performed by applying a transform from the left to the right, and then from the top to the bottom of the image block. Consequently, the top row and left column of the DCT coefficients are related directly to the detail in the corresponding row and column of the image block, but other coefficients do not have such a direct correspondence. However, the further away a coefficient is from the top left corner, the higher is the frequency of which it is a measure. In fact the coefficient in the top left-hand corner is the average grey level over all 64 pixels in the image block. In descriptions of the JPEG algorithm the top left-hand coefficient is called the DC coefficient, and all others are referred to as AC coefficients (this naming is derived from terminology used in electrical engineering in the theory of direct current and alternating current).

For natural images the size of the image block means that it will contain little in the way of detail, and hence there will only be low-frequency spatial components, with their nonzero DCT coefficients grouped around the DC component. The next step in the JPEG method is to quantize the coefficients but the number of quantizing steps is not the same in all cases. Some coefficients are quantized more coarsely then others, and, furthermore, the standard does not specify the detail of this quantization. For the decoding process, a table, specifying the quantization used for each coefficient, must be included in the file interchange format. Many of the coefficients after quantization are reduced to zero. It is

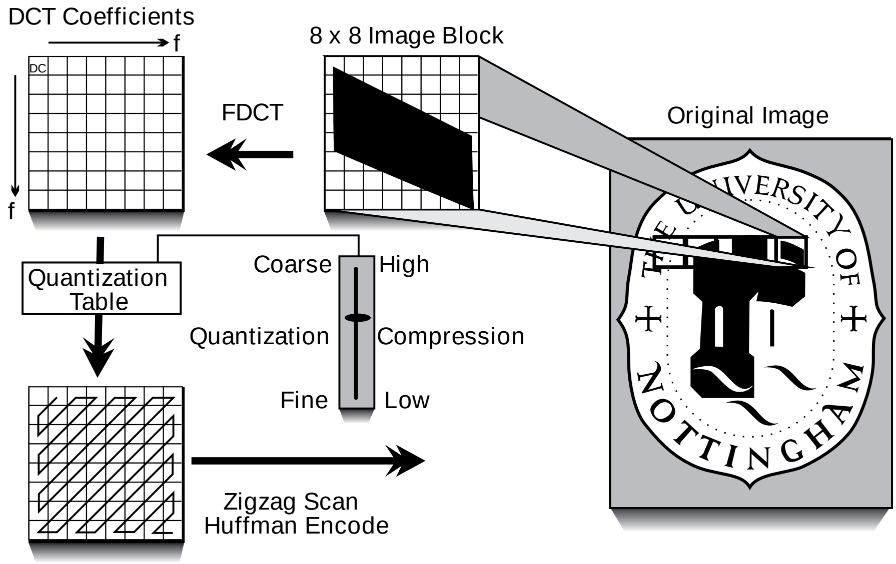

Figure 6. The basics of JPEG image compression

the specification of the quantizing levels, and the consequent reduction of coefficients to zero values, which gives rise to the compression in JPEG.

The last step is to encode the coefficients. Two methods of encoding are specified in the standard, one of which is covered by a patent, and which, consequently, is less frequently used. The other method is Huffman run-length encoding (see the section on fax encoding). Because the spatial frequencies in an image block tend to clump together, and as explained earlier the distance from the top left corner of the DCT array is a measure of frequency, the Huffman encoding is applied to the data stream obtained by traversing the DCT coefficients in the zigzag path shown in Figure 6. This path maximizes the zero-value run lengths, since the absence, or low amplitude, of a particular frequency in the image block will tend to give zero values to coefficients, after quantization, along the diagonals in the zigzag path. The Huffman encoding table is included in the output of the encoder for use by the decoder. A careful examination of the zigzag path in Figure 6 will reveal that the DC coefficient is not part of the path. This is because of the special nature of the DC coefficient and, in particular, that adjacent image blocks will generally have DC coefficients that exhibit only small differences from one another. Consequently, to improve encoding efficiency, the differences in the DC coefficients are stored using Differential PCM (DPCM).

The description of the algorithm just given assumes a single plane in the image. For colour there are three planes, and the algorithm could be applied three times. The decoder, when presented with such an encoded file, would be able to build up, sequentially, the red, green and blue components of an image (for example). This sequential build up, when decoding, is regarded as being unsatisfactory. To overcome this the encoding is done by interleaving values from the three planes, which, in turn, allows the decoder to build up the full-colour image line by line. The general practice for colour is to use one luminance and two colour difference signals (this colour space being the one used in colour TV). The luminance signal is the representation of the brightness of an image, made up of contributions from the red, green and blue components. The human visual system is more sensitive to changes in brightness than it is to changes in colour. This physiological effect allows colour information to be sent in a TV signal, or stored in a JPEG file, with reduced resolution for the colour difference components. A colour-difference signal is obtained by subtracting from the luminance signal the contribution made to the scene by one of the primary colours. The usual choice is to use red and blue colour-difference signals. Since the luminance signal is made up of contributions from the red, green and blue components of the scene, it is possible, with suitable manipulation, to recover the green component using the luminance together with the red and blue colour-difference signals. Note that, for every plane that is encoded, an associated set of tables must be included in the JFIF file. An example of JPEG compression of a grey scale image, with increasing compression ratio, is shown in Figure 7.

3.3.3 Fractal image compression

The Beauty of Fractals [18] is a book, which, amongst many other things, shows how fractal-generated images can take on a natural appearance. Many articles have shown how self-similarity in nature can be modelled using fractal-based algorithms, the classic example being the fronds of a fern. The breakthrough, as far as image compression is concerned, is in reverse-engineering the process. Rather than building natural-looking images using fractal-based algorithms, we want to take an actual image of the real world and derive fractal-based equations, from which the real image can be created.

The theoretical basis of this process is too extensive to include here, but Reference [19] is a good place to start if details of the theory are of interest; the exact nature of the current commercial exploitation of the theory remains proprietary information [20]. At the heart of the process is the fact that a set of transformation equations — which in two dimensions scale, rotate and shear one shape into another — will eventually produce a specific image if applied repeatedly to any arbitrary starting shape. In other words the equations, known as an iterated function system, IFS, are an encapsulation of the image. The arbitrary starting shape has an effect on the number of iterations needed to produce an output image of acceptable quality, and in practical implementations the image is broken down into small areas before an IFS is determined.

A major feature of the decompression method is that — unlike JPEG — it is output-resolution independent. This has to be taken in context, however, since the original image resolution cannot be improved upon. Nevertheless, if a decompressed image is viewed at a greater resolution than the original, the result is pleasing to the eye and shows no pixel duplication. Pushing the number of iterations further, in an attempt to see more than was originally present, will eventually produce artefacts not related to the original image. The time to compress an image is generally longer than that needed for decompression and maximum performance tends to call for specialized hardware. Software-only compression algorithms have been released recently and these are available in a number of ‘consumer-level’ software products. One such product has been used to produce Figure 8, which should be compared to Figure 7, which used JPEG, as an indication of the visual effect produced by these two techniques when comparable compression ratios

The Original Image 65,536 bytes

Compression 4.46:1 14,709 bytes

Compression 6.86:1 9550 bytes

Compression 17.49:1 3746 bytes

Compression 26.77:1 2448 bytes

Compression 38.44:1 1705 bytes

Figure 7. JPEG compressed images

The Original Image 65,536 bytes

Compression 4.59:1 14,290 bytes

Compression 7.40:1 8862 bytes

Compression 13.63:1 4807 bytes

Compression 29.19:1 2245 bytes

Compression 36.48:1 1796 bytes

Figure 8. Fractal compressed images

are in operation. In both figures the compression ratio is simply the ratio of the file sizes to the original bitmap size and it should be noted that the compressed file sizes include both data and header information.

3.4 Animation and video

In some respects animation and video are one and the same thing. The illusion of movement is derived from the same physiological effect — the only difference being that in animation the individual frames are drawn, by hand or computer, and in video the frames are captured from real life. In producing animation, however, there is one major advantage: the objects to be animated are well-defined. To use a fractal analogy, the equations are defined, so manipulation of the objects from frame to frame can be done by manipulating the equations, and regenerating the changed shapes of the objects. The only object available in video is the bitmap of the complete frame. To replay a sequence, whether animation or video, each frame in its entirety must be sequentially presented to the viewer. Storing a sequence of full-resolution bitmaps is far too costly in terms of data storage, so all major formats use compression techniques based on inter-frame correlation. Most animation formats are proprietary, although the file formats can usually be obtained. Proprietary video formats are currently fighting to gain market share, and detailed information on formats is only available by paying large sums of money. The only major non-proprietary video system is that from the Moving Picture Experts Group, MPEG-1 [21]. As might be expected from a standard which attempts to cover video and audio, there is no easy way to describe every aspect of the encoding/decoding process. Consequently, no attempt will be made to describe the audio aspects of MPEG-1, and for the video compression only the top-level methods will be presented.

In MPEG-1 two hierarchical video compression methods are in use, intra- and inter-frame. The intra-frame compression is very similar, but not identical, to the still-image compression of JPEG (see Section 3.3.2). The inter-frame compression makes use of the similarity between successive frames in a typical sequence of frames. Three types of frame are used, I, P and B. The I frames contain JPEG-like compressed images. When requiring random access to an MPEG-1 sequence, an I frame must be selected, and typically I frames are produced every one or two seconds for this purpose. P (predicted) frames are coded in such a way as to make use of previous I or P frames. Because a given P frame can use previous P frames there is a possibility of errors being propagated from frame to frame, but this effect is corrected at the next I frame that is encountered. Finally, B (bi-directional) frames are coded to be able to take information from the nearest I or P frames, past and future. There are no hard rules as to the relative position, or frequency of occurrence, of these three types of frame in an MPEG-1 sequence. In practice it is usual to have two I frames per second, and two B frames between every P frame at all other times. In order to make use of B frames, an MPEG-1 decoder must have available the future I or P frame upon which the B frame is based. Consequently decoders must be able to store a representation of a number of frames that are reconstructed out-of-sequence as far as the actual real-time output is concerned.

Motion compensation is one of the main data-reduction techniques used in the inter-frame compression. In this a block of pixels in one frame is identified as being almost, or exactly, identical to a block in a previous frame. Rather than encoding the block of pixels, the relative movement of the block is recorded, together with the difference (error), if

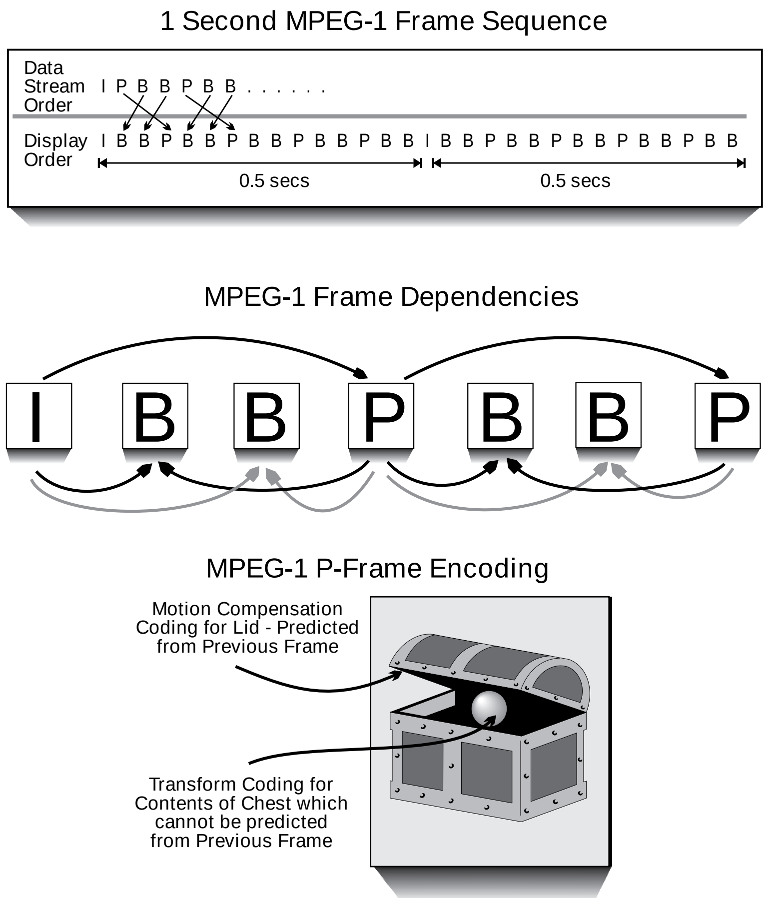

Figure 9. MPEG-1 fundamentals

any, between the two blocks. Figure 9 summarizes the detail presented here and explains why more than one encoding scheme may need to be used on P frames.

To produce pictures at a quality near or equal to that of VHS, the input data-stream rate, to an MPEG-1 decoder, has to match the output data rate of single-speed CD drives. A recent report [22] of a demonstration by a group of leading consumer electronics manufacturers described a system capable of playing back 74-minute video sequences from a standard-sized CD. At least one manufacturer [23] has shown that double-density CDs are possible, and so the possibility of full-length feature films on CDs is not far away. On 16 July 1993 an ISO/IEC press release announced the completion of the definition phase of MPEG-2. It is likely to be around 1996 before mass-produced implementations appear, but by then MPEG-2 may well be the basis for digital HDTV. The open nature of both MPEG-1 and MPEG-2 means that no single organization has a direct commercial stake in its adoption, and broadcast authorities, amongst other organizations, can wield tremendous influence in support of such a standard. Consequently, if MPEG becomes such a supported standard, there will be large investment in providing hardware support for a vast consumer market, which in turn means that the proprietary standards battling for position in multimedia will be overtaken.

4 CONCLUSIONS

In attempting to cover a broad range of topics this paper has touched on those of current significance, but it cannot be comprehensive. It has had a technological rather than an aesthetic bias, whereas a suitable synergy between these two has to be achieved by authors and publishers for any given publication.

The academic ideal is to apply a formalism to an activity by which the activity may be both studied and extended. Multimedia publishing presents a considerable challenge to a total formalism since it encompasses human cognition, high technology and commercial pressures. To date no single formalism has been successful in dealing with the simple printed word. Despite this, formalisms, such as HyTime [24], are beginning to emerge, which deal with the technology, but not the aesthetics, of multimedia. In the same way that good photography alone cannot guarantee a cinema success, so a formalism cannot guarantee a good multimedia product.

Electronic publishing, especially in the form of multimedia, is bandwidth hungry. The bandwidth is needed in two phases of the publishing process. Firstly, the transportation of the electronically encapsulated data, and secondly, internally, in the presentation of the data on a workstation or personal computer. Developments in networking promise much, but no-one can predict when the promised technology will be able to deliver acceptable bandwidth at the personal level. At present the penetration of ISDN into the business environment is small — and it is negligible at the domestic level. By using the Internet, academic institutions, including libraries, and an increasing number of commercial organizations, are beginning to have some experience of what a high-bandwidth, networked society might offer. Networking a society presents enormous cost associated with the upheaval of cable laying, no matter what the nature of the cable. One possibility for providing domestic-level access to networks, is to utilize the unused bandwidth of cable TV networks. PTTs may oppose this strongly, seeing it as a loss of revenue and a disincentive to their development of ISDN and B-ISDN. However, the electronics industry, in its short history, has always delivered the goods at the right price, when the market is sufficiently large, so the presentation side of the problem does not seem to be the limitation, if the demand is high enough.

In all of this discussion it must not be forgotten that the ubiquitous CD, and all its associated formats, offers an alternative dissemination system to high-bandwidth networks, particularly for information which is stable and long-term. Sending a box of ten, 650 Mbyte CDs, by 24-hour delivery service, is the equivalent, in data transfer rate terms, of having sole use of a 568 Kbps data link for 24 hours. The transit time is dreadful, but the packet size is gigantic ! Indeed, strategically placed CDs on a network offer a highly cost-effective way of allowing access to large amounts of data, and of minimizing network charges. Double- and quadruple-density recording, coupled with improvements in playback speed, make the CD an even more attractive proposition.

Interaction with CD-based packages offers new learning environments which are yet to be fully exploited though there is a great danger, in this area, of backing the wrong horse. The marketplace can be manipulated to allow technically inferior systems to gain supremacy — for example in the Betamax v. VHS battle to establish a standard VCR format. One encouraging development is the increasing speed at which certain open standards, such as MPEG, are being developed. This should mean that technical merit, and platform-independence, will play a greater part in determining the market success of technologies.

Authors and publishers will need to be adaptable in order to cope with the changes in electronic publishing. While single authors will still be able to produce enhanced, 'electronic', versions of existing material it may increasingly be the case that a team approach is needed in order to exploit all that the new technologies can offer. Publishers then take on the role of organizing the team, and investing in production facilities far removed from the printing press. Distribution of learned journals over networks is one obvious way that would greatly enhance the speed of dissemination. To do this, at present, a publisher would have to be an entity on the Internet, and also devise some method of deriving income from the distribution of the copyright material. Trying to predict the changes, whilst a pleasant way of passing the time, is like trying to predict the winner of a horse race. Past form is a good starting point, but the unexpected can, and usually does, happen.

REFERENCES

1. *Software Toolwork Multi-Media Encyclopedia*, 1992.
2. W. W. Wu and A. Livne, 'ISDN: A Snapshot', *Proceedings of the IEEE*, **79** (21), 103–111 (1991).
3. William Stallings, *Advances in ISDN and Broadband ISDN,* IEEE Computer Society Press, 1992.
4. R. W. Stephenson and S. A. McGaw, 'Southwestern Bell Telephone's ISDN experience', *IEEE Network Magazine* 25–36 (1989).
5. Ed Krol, *The Whole Internet: User's Guide and Catalog,* O'Reilly and Associates, 1993.
6. For information on public-domain MIME-compliant software, and the MIME specification do an anonymous ftp to `thumper@bellcore.com,[128.96.41.1]`.
7. Details of archie can be obtained from:

Host	IP address	Service area	Location
archie.doc.ic.ac.uk	[146.169.11.3]	UK/European	Imperial, London, UK
archie.funet.fi	[128.214.6.100]	European	FUnet, Helsinki, Finland
archie.au	[139.130.4.6]	Australian	Deakin, Geelong, Australia
archie.cs.huji.ac.il	[132.65.6.15]	Israel	Israel
archie.sura.net	[128.167.254.179]	World	SURAnet, Maryland, USA
archie.rutgers.edu	[128.6.18.15]	World	Rutgers, New Jersey, USA
archie.unl.edu	[129.93.1.14]	World	Lincoln, Nebraska, USA
archie.ans.net	[147.225.1.2]	World	ANS, New York, USA
archie.mcgill.ca	[132.206.2.3]	World	McGill, Montreal, Canada

8. For Information on Gopher and Gopher+ do an anonymous ftp to `boombox.micro.umn.edu` and look in directory `pub/gopher`.
9. *The Earth Dweller's Guide*, Available from Central Independent Television plc, East Midlands Television Centre, Nottingham, UK, first broadcast May 1992, repeated July/August 1993.
10. Philips IMS, Philips Electronics UK Ltd., *The CD-I Design Handbook,* Addison-Wesley, Reading, MA, 1992.
11. British Standards Institution, *Volume and file structure of compact read only optical disks (CD-ROM) for information interchange*, BS 7061:1989, ISO 9660-1988.
12. R. Ng, 'Fast computer memories', *IEEE Spectrum* 36–39 (1992).

13. Adobe Systems Inc., *PostScript Language Reference Manual,* Addison-Wesley, Reading, MA, 1990.
14. Adobe Systems Inc., *Adobe Type 1 Font Format,* Addison-Wesley, Reading, MA, 1990.
15. Adobe Systems Inc., *Portable Document Format Reference Manual,* Addison-Wesley, Reading, MA, 1993.
16. D. Bodson, K. R. McConnel and R. Schaphorst, *FAX: digital facsimile technology and applications,* Artech House, Boston, MA, 1992.
17. G. K. Wallace, 'The JPEG still-picture compression standard', *Comm. ACM*, **34** (4), 30–44 (1992).
18. H. O. Peitgen and P. H. Richterg,, *The Beauty of Fractals,* Springer Verlag, New York, 1986.
19. R Dettmer, 'Form and functions: fractal based image compression', *IEE Review* 323–327 (1992).
20. Iterated Systems Inc., Atlanta, GA. USA
21. ISO/IEC Draft Standard, *Coded representation of picture, audio and multimedia/hypermedia information CD 11172*, December 6 1991.
22. 'CD players with split personalities', *The Independent*, 14 (1993).
23. Nimbus Technology and Engineering, Wyaston Leys, Monmouth, NP5 3SR, UK.
24. L. A. Carr, D. W. Barron, H. C. Davis and W. Hall, 'Why use Hytime?', *Electronic Publishing — Origination, Dissemination and Design*, **7** (3), 163–178 (1994).