Introduction

Multimedia is the latest development in the abstraction of human intercourse that began with the stylized cave paintings of the earliest hunter-gatherers. These cave paintings were descriptors of daily life and say something of our primitive need to convey information (although to whom and why is not clear), which would lead to a codified, set of symbols we now call text. The earliest civilizations are notable by the change from pictures to pictograms and then to more abstract symbols as the need to record transactions both commercial and personal became increasingly important to the ordering of their societies. This can be seen in the development of the Sumerian cuneiform script beginning 5000 years ago (Cuneiform Writing, Undated). The development of writing on papyrus in Egypt, circa 4000 BC, (Dr. Ragab’s Papyrus Institute, n.d.) and leaves (Sri Lanka for example) allowed more cursive codes to be developed and more complex ideas to be disseminated to a wider audience. These materials were to be superceded in Europe with the introduction of pulp paper by the Arabs from Samarkand in 751 AD. It must be noted that the Chinese first invented paper. Attributed to Cao Lun in 105 A.D, there is evidence that it was known two hundred years earlier. Until the fifteenth century, scribes would copy texts to be read to the masses or for the education of a few. The above developments were slow, taking place over millennia, however the development of the printing press often credited to Johannes Gutenberg (Wikipedia 2008), a silver smith from Mainz, Germany in 1436 was to change everything. Within 60 years presses were available in 250 European cities. No longer would the printed word be in the hands of a few scribes and controlled by government bodies (notably in the West by the Catholic Church) but anyone could make copies of their work and, importantly, distribute them quickly. The Chinese were the developers of moveable type that is the key feature of the Gutenberg press. This is attributed to Bi Sheng, 1045, but their extensive character set did not allow it to develop in the way it did in the west. The days of mass media had not yet arrived but the means to achieve it had. It would take several centuries before presses were built using high-speed technologies and paper manufacture became cheap enough for a mass audience to be reached via newspapers. It was not long after the introduction of newspapers that pictures were appearing to illustrate articles. (The Illustrated London News, 1842, was the first fully illustrated newspaper). Initially these were often cartoons and drawings of the events but this soon gave way to photography. The Crimean War saw this change for the Illustrated Times newspaper in 1856. Its original commentaries and reports on the war were accompanied by illustrations sent from the front but by the end of the conflict photographs were being used. Throughout all this development images had supported text to an ever greater degree. It would require the advent of television to open the next channel for communication with pictures. Our increasing technological understanding led us to the development of networks and computers. Both originally designed with text in mind and both to become transmitters of images.

This module is about the collision of two technological worlds; Networks and PC's and two domains: the computer and the media, and their subsequent merging and evolution. It is only ten years ago that Multi-media became a serious proposition with the advent of cheap and quick enough PC's using a graphical interface to support a variety of media and it was in only in 1989 that Tim Berners-Lee first introduced a file exchange system which was to become the World Wide Web (WWW). That file transfer system was aimed at scientists and it was not until the first popular graphical browser (Mosaic by Marc Andreessen, 1993) came along that the WWW really took off. Yet the two; the Internet and Multimedia, have progressed to the point where they have been combined into a major international outlet providing leisure activities, education and business applications instantaneously across the world. That new world is the subject of this module; Internet and Multimedia Technologies. Indeed, it is the convergence of aesthetics and technologies.

We are all familiar with the Internet and its well-known application, the World Wide Web, but what is Multimedia? We might agree that it is not text alone but then neither is it purely graphical in nature. It can contain graphics, text, video and audio but only as required. It appears on a computer but only one that is capable of certain performance levels and contains certain minimum components. It may be simple or complex; it may be aimed at a few specialists or at a mass audience. So how are we to define it? Should we do that in terms of the technologies required or the presentational medium used? Several definitions have been suggested and here are two:

From the Encyclopedia Britannica, 2002: Interactive Multimedia; any computer delivered electronic system that allows the user to control, combine and manipulate different types of media, such as text, sound, video, computer graphics and animation.

The set text (Chapman) suggests:
Any combination of two or more media, represented in digital form, sufficiently well integrated to be presented via a single interface, or manipulated by a single computer program.

(I am confident you can think of more☺).

It is these definitions, which provide our stepping off point, and you will read more about its justification in this weeks reading assignment. Like all good scientists, it is our focus from which we examine our subject or (if you prefer) it is our mission statement, which constrains our business to that of the Internet and Multimedia technologies.
During this module we will begin by establishing the scope of our studies during this first seminar and progress to a detailed examination of some key aspects. What we cannot do is cover all topics in depth (if we did we would reduce our MSc to one module☺). It is divergent and like a tree produces branches for the interested to follow. Thus we will stick to the main body with occasional forays along the branches to explore its potential.

This week, after the introductory discussion, we need to briefly visit the enabling technologies. How do we digitize our media? What hardware and software do we need and when? How can we distribute these ideas? What standards can we call on to ensure we are understood? Each of these will be expanded on later as required but for now let us begin with digitization.

For the mathematicians amongst us, we could begin with a discussion of the Dirac Delta function (Bock, R.K, Krischer W. 1999) (it can be thought of as a function $\delta(x)$ that has the value of infinity for $x = 0$, the value zero elsewhere, and a total integral of 1) which allows us to consider operations at a point and justify our digitization process in a rigorous manner but this is an introduction and we have too great a distance to go in too short a time. You may wish to refer to any good book on image processing for a detailed discussion, e.g. Digital Image Processing by W.K. Pratt. So let us look at digitization in a more pragmatic manner. We need ultimately to translate all our work and all the input media to a digital form for the simple reason that computers only understand a 1 or 0, off or on, state. Everything we do must, therefore, be coded in a binary form for digestion and transmission by a computer. The question is simply, how can we achieve this for different media? One approach is to look at the different types of media and see what their nature is. We must consider four key types; text, graphics, audio and video.

If we take text first, then we have a code whose elements are our respective alphabets. Superficially, this one is easy; we simply assign a binary number to each element. For many years we in the computing industry focused on the Roman alphabet and its punctuation system. This would fit easily in a seven bit code (known as the American Standard Code for Information Interchange, or ASCII, code). Since there are eight bits in a byte, this was convenient and allowed the extra bit to be used for simple error checking. Unfortunately, it was soon evident that this could not accommodate even the European languages and so many variants sprang up using an eighth bit. Finally, with increasing Internationalisation, it was accepted that this was insufficient and so the 16 bit Unicode (Unicode Home Page, Anon) came into being. It is this code which allows the development of applications which can be targeted across language and platform barriers worldwide. We will return to the subject of text in Seminar 4. Text then, is a relatively simple form to digitize but what about graphics?

For day to day purposes, an image is made up of continuously varying intensities of light and similarly varying colours. A ‘graphic’ is an interpretation of this fundamental state. When we take a photograph using film, we limit the fidelity in a variety of ways. The range of colours is reduced and the light intensity is restricted in range however we have a sense of continuity because of the random nature of the silver halide particles on the film. With a digital camera, we assign a pixel (an array element) to each discernable point in our image. Typical values are 640 x 480 pixel array. Into each element we must assign a value which represents the light intensity and colour at that point. To do this we provide a range of binary values to represent the light intensity and another range to represent a chosen colour. In complex 3D graphics additional values or attributes are required which describe how to modify the element during rendering (the process of transferring the information to the screen). Thus for each element or pixel, we may use 32 bits or more for high fidelity work. This is not typical of web based applications. How we do this, we will look at in detail in seminar 2. Now we can store each pixel description in a computer file. If we know the name of the file and how it is structured we can display the graphic any where in the world where there is a suitable computer. Compared to text our average file size has increased considerably.

Let us look at audio next. (It is clearly developed in more detail in the text). Audio begins as a set of continuous frequencies and for each one we must assign a value which describes its amplitude. However that is not sufficient. We need to know precisely where in time each occurs. There are many ways of achieving this but the following should make the process clear. In humans, sound is generated by the lungs vibrating the vocal chords. A range of frequencies at different amplitudes are produced and therefore analogue devices (such as a tape cassette) are suitable for recording and reproducing these sounds. However they do not suite our digital technology. We need to change the analogue values to digital for storage and then return them to analogue for the listener. To further complicate the digitization process we need to understand that human sounds may be considered as two groups ‘voiced’ and ‘unvoiced’. These correspond to high amplitude and low amplitude values respectively. Most of the power of human speech is carried in the voiced sounds but we differentiate between similar words using the unvoiced sounds. We will see later that silence is also a key element of conversation and speech. We need to account for all these features if our digital storage and transmission is to be efficient for audio. .

We have ignored music and its greater frequency range for the moment. For the moment, let us translate the human voice to digital. Alexander Bell realized that the full range of frequencies of the human voice, approximately 60 to 11000 Hertz, (A Hertz is a frequency of one cycle per second) were not required for transmission on the telephone. He discovered that only a small acoustical power was carried above 4 kHz and that, in fact, 80% of the power lies in the frequencies 300Hz to 3300Hz. (This was to have severe repercussions when we later required high fidelity for music). Thus industry developed systems that worked on channels spaced at 4 kHz intervals.

Digital data has several additional advantages over Analogue data. The fact that only specific signal levels are valid implies that the random noise that might be injected into the analogue signal and thus distorting it (because they are undistinguishable from every other signal), does not interfere with the digital signals. That is because with the threshold surrounding the digital signal all is considered to be the specific signal, while if the value does change to a non-valid signal it can be restored to the original signal. Even distortion due to high values of noise can be detected due to various error detection and correction routines. Thus the digital information consists of a much more noiseless channel (in the sense of Shannon's Information Theory (Shannon 1948)) than the analogue one.

So how do we go about the digitization process?
There are three steps; sample, quantize and code.

We begin with an analogue signal that we wish to reproduce as a digital code. We first need to choose a set of times to sample the analogue signal amplitude. If this time is too short we will take unnecessary samples and if it is too long we will not be able to reproduce the signal accurately. Fortunately, there is a known rule which allows us to set the sampling rate appropriately.

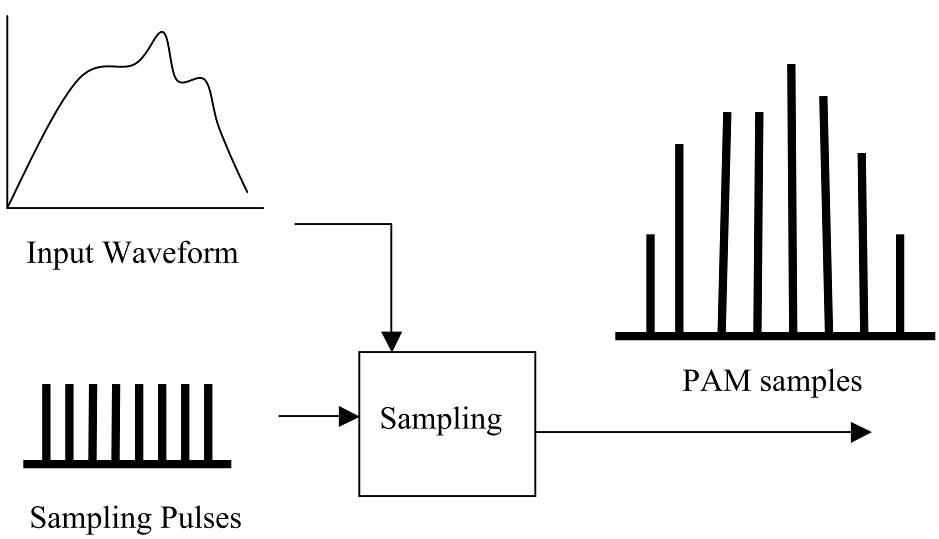

Figure 1. PAM sampling.

This is the Nyquist Sampling rate and was developed by H. Nyquist in 1933 (Roland-Mieszkowski & Young 1991). It states that the minimum rate needed is twice the highest frequency we wish to preserve. E.g. For 4 kHz, we should use 8000 samples per second. Sampling produces a set of Pulse Amplitude Modulated, PAM, samples that are pulses which follow the amplitude curve of the analogue signal. See figure 1.

We must now give each sample a value which represents its amplitude. This is quantization. It was decided that each sample would be represented by an eight bit number and therefore we must transmit our signal 8 x number of samples or at 64Kbits per second. The computer must therefore store 64kbits of data for every second of speech without compression. This is not the end of the story for networks. The final stage is to encode the signal stream so that it can be transmitted efficiently and without error. We will discuss this in seminar 8.

I will leave you to read chapter 2 of the set book and examine its clear illustrations on this topic however be aware that this is a complex subject from which, we only require the essentials.

Music is a more complex problem requiring greater bandwidth. When an instrument vibrates it produces sound waves of a particular frequency. However, many instruments can produce the same fundamental frequency. What sets each apart is the production, amplitude and number of multiples of the fundamental frequency that the instrument produces. These waves are called harmonics.
The basic frequency and its harmonics determine the quality or timbre of a sound. The greater the number of harmonics, the more interesting is the sound that is produced. It is an object's ability to vibrate and set up harmonics that determines the pleasantness of the resultant sound. For example, crystal glass produces harmonics that are more pleasant than those of ordinary glass.

This combination of fundamental frequency and its harmonics produces a complex wave form and many terms have been developed to describe their form. Three of these, attack, sustain and decay collectively describe the envelope of a sound. See Figure 2.

From figure 1 we can see sound begins at A and reaches a peak volume at B. This is the Attack. Between B and C the sound continues and in our example is dropping in volume. This is Sustain. Finally, between C and D the sound diminishes until there is silence. This is the decay.

It is this envelope that we need to recreate accurately. Since some of us can hear even up to 20 kHz (when we are young) a much larger bandwidth is required to satisfactorily digitize music. A CD uses 44.1 kHz.

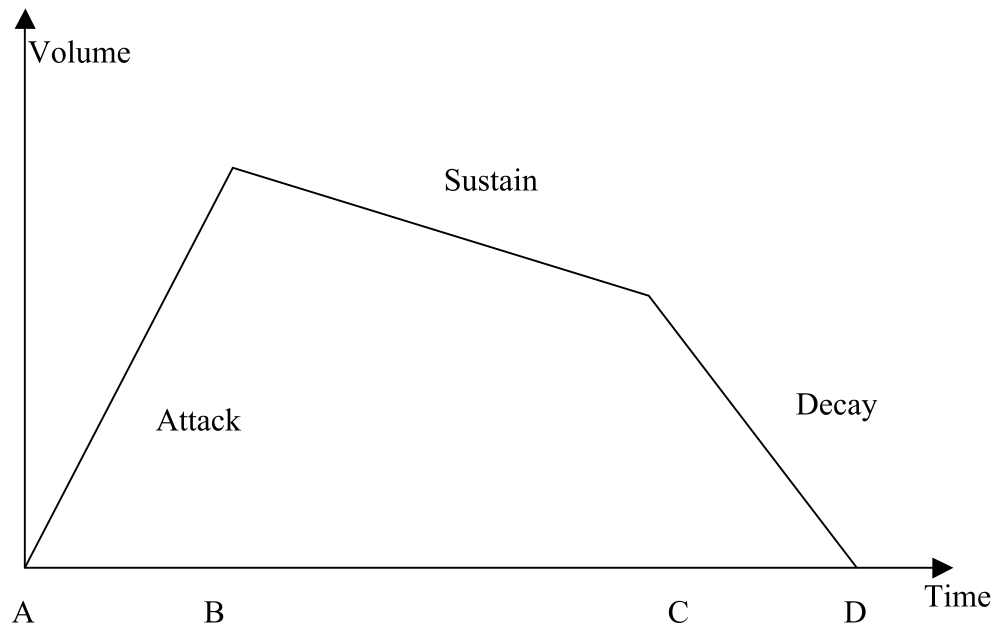

Figure 2. The Envelope of Sound

Finally, we have video. A video, similar to a film movie, is a collection of still graphics played back at a rate which convinces our brains that it is continuous view we are receiving. This depends on a number of parameters that include target size, adaptation level, and the eye's integration time. This varies slightly from person to person. Then depending on our purposes we must set the number of graphics (their frequency) to an acceptable frequency. For the television system we use 50 or 60 Hertz. (30 Hertz for systems based on the French SECAM technology). However for animations we consider 30 Hz's adequate. Films use 24 frames per second but the shutter works at 24 thus giving us the illusion required. These choices have considerable effects on the storage requirements for a given video. If we then add a sound track, which must be synchronized to the frames of the film, then our storage problem is considerably greater.

But what do we mean by storage? For our ancestors, it was beads and then clay tablets leading to paper. Computers were once heralded as the replacement for paper. Although we can all smile at that one currently, it would not be possible to produce everything we store on computers as paper copies today. Storage is appropriate to the time and place. The need for it has grown as our technology has advanced and our demands increased. We still find CD's adequate for most needs including Multimedia however we require DVD for films. On a personal note I am already moving everything to DVD for archiving and there are already new developments of even greater storage capacities (Shim 2003).

This is permanent storage, hard copy, the new age paper, … but accessing that data is comparatively slow; that is why we need local storage in the form of RAM with its higher access rates to feed the screen and the output devices in a timely manner. We will return to this issue at intervals in our programme but for now let us consider some of the ramifications for development.

Currently computer speed is doubling every 18 months whereas Network speed doubles every nine months. Does this imply that Networks are faster than computers? No, of course not. The two items are separate and the answer depends on what aspect you are measuring and the devices compared, however it does demonstrate that the interface adapter speed will increase and therefore our use of bandwidth intensive forms of Multimedia (perhaps including virtual reality) will also increase. The true test for the network is capacity; how much can an individual use and at what rate. By improving capacity, the individual can approach true broadband bandwidth (Assuming broadband to be the FCC definition of 200kbps). Unfortunately, counter balancing this is the fact that more users will want greater bandwidth and hence greater capacity will be required. This is a typical development cycle .See Figure 3:

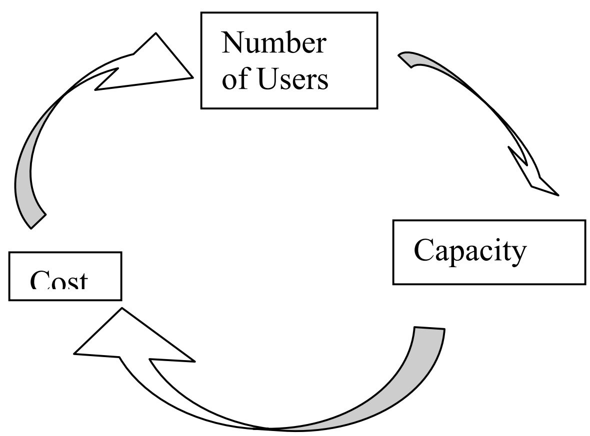

Figure 3. The Network development cycle

We may want greater bandwidth, but can we afford it? Will an increase in numbers of users reduce the cost to the extent that the network suppliers will need to provide more infrastructure? Can they afford to install new infrastructure? Will the government allow them to install it?

A similar problem exists for the computer developer. We could look at a computer system as a mass storage device, a processor and associated memory (RAM) and a display.

An increase in picture fidelity requires an increase in the number of pixels displayed so that manufacturers are encouraged to develop bigger and better displays. To display these pixels requires faster memory access and larger local memory; to fill this memory requires faster processors and access to faster mass storage which in turn needs to be larger to store all the additional information. Developments in any one area allow another area to change leading to increased development costs and for the user's equipment, obsolescence rather than component failure.

Not so long ago, network capacity was insufficient for a satisfactory multimedia experience and distribution was usually on CD's. (This is still the case for some products and back up facilities). Today the Multimedia companies offer downloading of their software as standard and multimedia is commonplace on the Internet. To see how this is achieved requires an understanding of the client-server model. Simplistically, a server is a computer which responds to request from its clients. In reality, a server may comprise several computers in different locations responding to many clients. Most of us as home users are in the position of the client making a request for service. Special network software on the server recognizes the requests and distributes them to the appropriate application which then formulates the response.

This raises several questions; how does the server and client know each others address, how does the software deal with different platforms and operating system environments, is it all standardized?

It is the World Wide Web which allows most interactivity today (there are, of course alternatives). It uses the Hypertext Transfer Protocol, (HTTP) to provide document exchange. These documents are written in the Hypertext Markup Language, (HTML), (More on that in Seminar 4). Please note that the initial idea for a linked structure was introduced many years before; in fact in 1945. The seminal paper by Vannevar Bush can be accessed via reference (Bush 1945). The client and the server both understand the notion of a Uniform Resource Locator (URL) which provides each side with an address. This is the well known web address which takes the form *protocol://servername.domainname/file locator/file.type* e.g. http://www.upenn.edu/museum/Games/cuneiform.shtml. From this we can see that the server is a web server using the http protocol with a computer name www located in the domain name, upenn.edu and the rest tells the server to locate the file cuneiform.htm. At the client end we use a browser to view the files e.g. Internet Explorer, Netscape etc. The browser has built-in functionality which allows it to respond to different file types in an appropriate manner. This information is built into the file header and works through extensions to the Internet Mail Protocol known as the Multipurpose Internet Mail Extension, MIME. You can see this in Gmail. if you open a message with an attachment, there is a scroll-down menu right off the "Reply" tag with the option "Show original". This will open a new window with the text version of the message and the content descriptors will be visible. This can also be done in Outlook by saving emails or webpages from IE to "Web Archive, single file" mht format. When you open the file using a text editor, the first few lines are the content descriptors**.**

None of this would be possible without standards that can be applied universally by all providers of software. Today there are many international standards

governing all sorts of technologies. These are brought together by the International Organization for Standards (ISO). Currently the demand for wireless applications is creating a considerable number of variations on existing protocols and several new ones. These applications use a variety of multi-media approaches to file interchange that have become increasingly popular as web services. Originally the standards were driven by businesses requiring data exchange. Today businesses have emerged which are solely dependent on the presence of exchangeable data (Multimedia businesses in particular) and it is these which now demand new standards for interoperability.
The Internet is extraordinary in the sense that it began as a collection of voluntary and, mostly, open bodies collectively determining the future of the Internet in the world today. The Internet Society, ISOC, which was established in 1992 consists of professional Internet experts and concerns itself with policies. It oversees a variety of boards and task forces dealing with Internet issues.

Several key associated groupings exist such as the Internet Engineering Steering Group, IESG, the Internet Architecture Board, IAB, the Internet Engineering Task Force, IETF, and the Internet Assigned Numbers Authority, IANA. The IETF is the protocol engineering and development arm of the Internet. The group was formally established by the IAB in 1986 (it had existed informally before that). It is this group which proposes protocols for consideration as standards although it is possible for any organisation to do so. This is due to the use of the Requests For Comments (RFC) process. The technical management of IETF activities is conducted by the IESG. It also has responsibility for the Internet standards process. The IAB, originally known as the Internet Activities Board, first came into being in 1983. It was behind the organised approach to the Internet. The IAB formed the IETF and the IRTF (Internet Research Task Force) in 1986. In 1992, it was re-constituted as the Internet Architecture Board and now serves as the technology advisory group to the Internet Society. It is responsible for defining the architecture of the Internet. IANA is in charge of all IP (Internet Protocol) addresses and any other parameters which are defined for the Internet. A major part of its task is to ensure the uniqueness of the parameters. There are numerous other standards and we will meet some of them in the course of the programme.

We have had a brief introductory look at the enabling technologies and we will return to them as and when appropriate but what of the software which allows us to produce our multimedia presentations?

There are many choices of software and there was a time when packages had a distinct functionality but as the major players developed their offerings began to overlap considerably. Some traditional favourites are the Macromedia products; Flash Mx, Director, breeze, Coldfusion etc., which handle most types of multimedia product from e-commerce to teaching. For 3d Graphics and animation work, 3D studio Max is probably the most popular commercial package. There are many others and designers develop their own favourites based mostly on look and feel of the tools once their basic needs are met. For

image manipulation Photoshop meets most people's needs.
The new multimedia world (and we should mention Virtual Reality as well) is indeed what Marshall McLuhan described as "The Medium is the Massage. (McLuhan 2005)": Society has always been "shaped more by the nature of the media by which men communicate than by the content of the communication"

An excellent "web book" with a lot of information on sight and graphics and a large number of demos and images is to be found in "The Joy of Visual Perception " (Kaiser 1996).

Additional Reading:

(Not a 'must do'☺)

If the topics covered appeal to you and you are interested in increasing your understanding, the following books should be of interest:

Cotton B. & Oliver R. (1997) Understanding Hypermedia, Phaidon Press, London

Crowcroft J.,Handley M. & Wakeman I. (1999), Internetworking Multimedia. ISBN 1-55860-584-3 Morgan Kaufman (Available via Google books: http://books.google.com/books?hl=en&id=xJs9AAAAIAAJ&dq=crowcroft+handley+wakeman&printsec=frontcover&source=web&ots=1MIXtWhCjv&sig=tq-BhvTx7KNw8CP-q0dfWoPfwJA&sa=X&oi=book_result&resnum=1&ct=result)

Bidgoli H. ed. (2003), The Internet Encyclopedia ISBN 0-471-22204-6 John Wiley
This is a very comprehensive text containing 250 articles. It is also very expensive. Hopefully your library will have a copy. Definitely not a recommended purchase unless you are extremely well – off ☺

McLuhan M. & Fiore Q. (2005) – The medium is the massage. An Inventory of Effects, (Gingkopress)

Pratt W.K. (2001), Digital Image Processing: PIKS Inside, Wiley-Interscience; 3 edition (July 24, 2001)

Links and References

Bock, R.K. & Krischer W., eds. (1999) The Data Analysis BriefBook [Internet]
http://rkb.home.cern.ch/rkb/AN16pp/node57.html#SECTION000570000000000000000
A brief description and explanation of the Dirac Delta function
(Accessed: 15th September 2008)

Bush V. (1945), As We May Think
The Atlantic Monthly, July 1945. Available from;

http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm,
A look at the future of meaningful human data retrieval
(Accessed: 15th September 2008)

Cuneiform Writing, Anon A short history of cuneiform [Internet]. Available from:
http://www.upenn.edu/museum/Games/cuneiform.html
(Accessed: 15th September 2008)

Dr. Ragab's Papyrus Institute, Anon [Internet]
http://www.papyrusinstitute.com/
(Accessed: 15th September 2008)
The story of the re-discovery of Papyrus technology by Dr. H. Ragab and a discussion of early Egyptian writing.

Kaiser P.K. (1996), The Joy of Visual Perception: A Web Book [Internet]
Available from; http://www.yorku.ca/eye/
(Accessed: 15th September 2008)

McLuhan M. (2005), The Medium is the Massage, Gingko Press
http://www.marshallmcluhan.com/
(Accessed: 15th September 2008)
The official site of Marshall McLuhan contains links to many of his works

Roland-Mieszkowski M. & Young W.R. (1991) Proceedings from "Acoustic Week in Canada 1991" - CAA Conference, Edmonton, Alberta, Canada, October 7 - 10, 1991
Available at http://www.digital-recordings.com/publ/pubneq.html
(Accessed: 15th September 2008)
A technical paper examining some of the ramifications of the Nyquist formula for Acoustics.

Shannon C.E. (1948) A Mathematical Theory of Communication by Claude E.Shannon
Reproduced in http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html
(Accessed: 15th September 2008)
Claude Shannon's "A mathematical theory of communication"

Shim R. (2003) Toshiba, NEC see blue in DVD future, CNET News.com
Available from http://zdnet.com.com/2100-1103_2-1001033.html
(Accessed: 15th September 2008)
One of the future proposals for the DVD

Unicode Home Page [Internet]
http://www.unicode.org/
(Accessed: 15th September 2008)

A detailed description of Unicode and its uses

Wikipedia (2008), Johannes Gutenberg [Internet]
http://en.wikipedia.org/wiki/Johannes_Gutenberg'
(Accessed: 15th September 2008)