Analog vs. Digital

"An important part of audiovisual archiving is the conversion of legacy audiovisual content from an analog to a digital form. "

Information Representation

In this discussion about audiovisual content, the term *information* will be used as a generic replacement for the sounds and images contained in audiovisual assets. In order for this information to be either conveyed from one location to another or preserved for later reproduction, a technical system of some form must be employed. These systems can be chemical, electrical, mechanical, or most typically, a combination of two or more of these technologies. This is true whether the system is analog or digital. The difference between these two classes of systems is the method by which information is represented.

Digital Information Representation

In a digital system, information is represented by a set of numbers. Each number in the set describes a specific aspect of the information at a specific instant of time. As a whole, the set describes all aspects of the information over a period of time.

In order to convey digitally represented information from one location to another, the set of numbers must be transported between the two locations. This can be done using one of several methods, such as via the Internet or over local area networks. In order to preserve digital information for later reproduction, the set of numbers must be stored. Sets of numbers may be preserved in a variety of ways, for example on computer discs or in permanent digital memory circuits.

Once they are recorded, the method by which the set of numbers is interpreted must be available so the audiovisual content can be reproduced. These methods are primarily embodied in software algorithms, which are typically executed by general purpose digital hardware capable of performing a variety of functions. In some cases this hardware may need special-purpose components which are added specifically to help reproduce the audiovisual content, although the need for such additional components has dropped as they are increasingly integrated into general purpose hardware.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

Analog Information Representation

By contrast, analog systems do not convey or preserve numbers, but represent information in entirely different forms. Electrical analog systems represent each aspect of information as a varying electrical quantity, such as voltage. This quantity varies continuously over a period of time. These varying electrical quantities are referred to as *signals*. Chemical analog systems, i.e. film, represent information as varying chemical densities.

It should be noted that while mechanical technology is vitally important to both analog and digital systems, it is not central to the representation of audiovisual information. Furthermore, since electrical and chemical systems are radically different, this discussion will focus on the contrasts between electrical analog technology and digital technology.

Electrical analog systems convey information between two locations using several means, most commonly through wires or over air waves. Information can be stored for later reproduction using a variety of methods, which can include cutting meandering grooves on a phonograph or orienting iron particles on a magnetic tape. The methods used to interpret the conveyed or stored signals are primarily embodied in hardware devices (often called players or receivers) which are dedicated to reproducing the audiovisual content.

Information Preservation

Whether in an analog or digital form, the audiovisual archivist's task is to preserve audiovisual content in such a way that it can be reproduced in the future. This means the content not only has to be preserved, but also the preservation has to be done in such a way that the content can be reproduced in the future.

There are two activities in which the archivist must engage in order to preserve audiovisual content. First, the archivist must ensure careful storage of the information. The goal of this is to preserve the information in such a way that the audiovisual content is not affected by any form of deterioration over time. This generally requires the archivist to provide a suitable stable environment in which the information is preserved. When preserving content in an analog form, this usually means that the media used to store the information, such a phonograph or magnetic tape, will not undergo physical changes over time. Preservation of content in a digital form may also require the archivist to provide a suitable stable environment for some form of storage media such as computer discs, or it may require continuous operation of an active digital storage system.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

Providing such an environment over long periods of time at a given location often does not provide enough certainty that no loss or alteration of the stored content will occur. Fires, floods, war, and other severe catastrophes have been known to cause the destruction of valuable information and artifacts throughout history. Likewise, there has never been an active digital storage system that did not eventually fail. In order to protect against such loss, it's common for archivists to store content in more than one location or in more than one system. When storing content in more than one location or system, it becomes necessary to have two or more copies of the content. Making exact replicas of information is therefore the second activity linked to the challenge of preserving audiovisual content.

Making Copies

It is much more difficult to make an exact replica of information in an analog form than information in a digital form. In fact, it's essentially impossible to make an exact replica of information in an analog form. The phrase *generation loss* comes from the inevitable alterations that occur when a copy is made of information represented in analog form. On the other hand, digital information, i.e. sets of numbers, can easily be replicated. Furthermore, it's usually a simple matter to confirm that a copy of digital information is a perfect replica of the original.

This distinction between analog and digital replication exists largely because the world we live in is essentially an analog environment. The brightness of light or the loudness of sound is not experienced as a set of numbers, but as continuously variable attributes. Devices which have been invented to establish the brightness of light or the loudness of sound simply convert these physical attributes into electrical quantities that vary in step with the measured attributes; analog systems convey or preserve these varying electrical quantities.

Digital systems, on the other hand, examine each varying electrical quantity produced by these measurement devices, instant by instant, and produce a number corresponding to each quantity at each instant of time. This conversion of varying quantities into sets of numbers isolates the digital information from the physical world by adding a layer of abstraction. This isolation of information from the physical realm forms the basis of a digital system's ability to produce an exact replica. Once isolated, the information itself becomes immunized from influences in the physical world that would act to change it.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

As an example of this distinction, think of a list of 100 numbers written in pencil on a piece of paper. Assume that this list represents the loudness of sound over a brief period of time. Each number in the list represents the loudness at each instant as the loudness varies and time progresses. To make an exact replica, an archivist simply copies the list of numbers onto another piece of paper. This archivist would not be hindered if the pencil used to make the original list was dull and the writing was slightly smudged, and even if the paper was torn and taped back together the archivist making a copy would still have a very good chance of making an exact replica. As long as the physical defects present in the original list are not overly severe, an exact copy of the list of numbers can be reliably produced and the information represented by those numbers perfectly replicated. The information, being in a digital form, is isolated from, and thus immunized against the influence of physical defects.

Now think instead of a meandering horizontal line drawn across a piece of paper. The line is drawn so that the distance between a point on the line and the bottom edge of the paper represents the loudness of sound at an instant of time. This line will rise and fall as it continues across the paper, just as the loudness varies and time progresses. This line is an analog equivalent to the list of numbers as its continuously varying distance from the bottom of the paper represents the changing loudness. While the numbers representing digital information in the previous example were easily copied, consider the challenges the archivist endeavoring to make a copy of this line on another piece of paper faces. Even if the archivist painstakingly measures the distance between the line and the bottom of the page in several places, or traces over the line on another piece of paper, any reduction in the quality of the original makes it impossible to perfectly reproduce. If the original line was drawn with a dull pencil and is a bit fuzzy, or if the line has become slightly smudged, the distance between a point on the line and the bottom of the paper becomes impossible to accurately determine. It may even be the case that the paper has shrunk, unevenly changing the distance between the line and the bottom of the paper. This archivist can successfully copy the line, but there is almost no way for the archivist to perfectly copy the line. Any physical defect in the original, no matter how minor, makes it essentially impossible to exactly reproduce. Any analog system, electrical or otherwise, is subject to physical influences that alter the represented information.

Thus, information in digital form is much more readily replicated than information in analog form. Furthermore, mathematical computations known as *checksums* provide a means for any variations between two lists of numbers to be reliably detected. When checksums are used together with other techniques, errors can even be corrected. Methods exist for measuring the degree of variation between two analog copies, but analog methods cannot recognize or recover the original information.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

Even the act of reproducing content stored as analog information will, in many cases, permanently alter its represented information. For example, every time a phonograph is played, its grooves are worn down and slightly changed. This will alter the sound reproduced when the record is played and the original sound is forever lost. The difficulty in producing exact analog copies is one reason why archivists find keeping audiovisual information in a digital form so attractive.

Digital Data

Digital systems, however, are not always ideal. It takes a huge set of numbers to accurately represent variable quantities. This fact was a major stumbling block to the development of digital audiovisual systems. Moving images in particular are extremely demanding of digital systems. Only with the improvement of digital technologies has the representation of audiovisual content become possible. Even still, departures from ideal digital systems are necessary in order to make these systems practical.

An important distinction between analog and digital systems is the process used by digital systems known as *sampling*. As previously mentioned, each number in the set of numbers describing audiovisual information, represents an aspect of that information at an instant of time. The process of making periodic measurements of a continuously changing quantity is called sampling. For sampling to produce a set of numbers which accurately describes the changing quantity, the measurements must be taken frequently. The frequency at which the samples are taken is called the *sample rate*.

Each measurement (or sample) results in a number. This number lies somewhere along a scale that spans the possible range of variation in the quantity under measure, from minimum to maximum. That scale divides the span of variation into divisions or steps. The smaller the steps, the more precise the measurement can be. For example, the Celsius scale divides temperature measurements between freezing and boiling water into 100 divisions, called degrees. Thus, in the Celsius scale, there are 100 degrees between freezing water, at 0 degrees, and boiling water, at 100 degrees. The Fahrenheit scale divides the same temperature range into 180 degrees between 32 degrees at the freezing point and 212 degrees at the boiling point. The Fahrenheit scale can therefore provide a more precise temperature measurement than the Celsius scale because it offers nearly twice the divisions over the same temperature span. A more precise scale requires a larger range of numbers, which typically requires more digits. Digital systems use the *Binary* numbering system in which each digit is referred to as a *bit*. The number of bits used for each sample is called the *bit depth* and this determines the precision of each sample.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

The sample rate and bit depth together determine the rate at which the numerical information (or *data*) representing audiovisual information is generated. This rate is called the *data rate* and is a good measure of how much demand is placed on the capacity of a digital system. Other factors affect data rate as well, such as stereo vs. mono or black & white vs. color. Sample rates and bit depths are largely standardized such that the numbers used by digital systems accurately describe the varying quantities they represent. Since the standardized data rates generate large amounts of data very quickly, especially when used to represent moving color images, these data rates can overwhelm even modern digital systems.

Digital Compromises

In order for digital systems to more easily convey and store digital audiovisual information, a process known as *data rate reduction* is used, which is also referred to as *compression*. This process endeavors to reduce the data rate to a level that digital systems can accommodate without producing a perceivable change in the audiovisual information. The ratio of the original data rate to the reduced data rate, known as the *compression ratio*, along with the effectiveness of the methods used to achieve it, affect the performance of many digital audiovisual systems more than any other factor.

It's not unusual for compression to be intentionally applied to such an extent that information degradation becomes apparent to a sensitive observer. This is done to allow a low capacity digital system to convey or store audiovisual information. The Internet is a good example of a low capacity system. While the Internet's capacity has increased dramatically since its inception, very high compression ratios are required to support real-time audiovisual content delivery over the Internet. Cell phones are another good example; while cell phone memory capacity has increased over time, it would not be possible to store hours of music on a phone without employing high compression ratios. In both these cases, the benefits of using high compression ratios are considered to be of such value that small amounts of perceivable information distortion is accepted as a satisfactory trade-off.

Compression can be tailored to match a wide variety of digital systems. Expensive high performance digital systems exist which have the capacity for audiovisual information without the need for any compression. Other systems exist with enough capacity to accommodate audiovisual information to which *lossless compression* has been applied. Lossless compression is a technique that produces a modest decrease in the data rate such that no information is lost. This is done simply by removing data that is redundant or, to put it another way, can be expressed more efficiently.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

The next step along the compression spectrum results in some loss of information, but that loss is unperceivable. Research has been remarkably successful in identifying information that can be removed without our eyes or ears perceiving the loss. This type of compression utilizes extremely complex computational techniques that rely heavily on digital technology. These techniques can achieve greater compression ratios with only minor and unnoticeable information loss.

In order to accommodate low performance digital systems, these same techniques can be applied more aggressively, reducing the data rate even further, but with a noticeable loss of information. The degree to which the loss is objectionable depends on the observer. Compression applied to cable TV and streaming video content falls into this compression range and is unnoticed by most viewers, although some find it objectionable.

Building Tailored Digital Representations

The processes of sampling, compressing, and organizing audiovisual data are collectively referred to as *encoding*. If audiovisual information has been encoded with minimal compression, it can be re-encoded (or *transcoded*) using a higher compression ratio and thus made available to lower capacity digital systems. Archivists can therefore take audiovisual data with little or no compression and create a variety of copies tailored to different digital systems. The process of compressing data can be reversed however there are few occasions to decompress data for archival purposes because the information lost during *lossy compression* cannot be recovered.

Reproducing Content

The audiovisual archivist must also address the challenge of reproducing audiovisual content. Even if audiovisual information has been successfully preserved for future use, it can only be reproduced if the technology and knowledge needed to do so is still available. Indeed, the lack of available functional analog players and knowledgeable operators is a driving force behind the migration of audiovisual information from analog to digital. This does not entirely solve the problem, however, as the rate of change for digital technology is greater than it ever was for analog.

In order for digital audiovisual content to be reproduced, the encoded data must first be *decoded*. The organization of the data must be understood so that the meaning of each number in the set can be determined. The method of compression must also be understood so the compressed data can be expanded.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446

The ability to decode audiovisual data is embodied in instructions or *software*. This software is essential to the process of reproducing preserved digital audiovisual content; however, software alone is not all that's needed. In addition, *hardware* must be available that can recover data from a storage medium and execute the instructions for decoding the data. If preserved content is to be reproduced, both the needed software and hardware must be available.

Choosing the Right Technology

One approach to selecting a specific digital technology for archiving audiovisual content is to select the highest performing technology. This assumes that the technology will remain available in the future by virtue of its high performance. Another approach is to select nonproprietary technology. In this case, it is assumed that since the technology is in the public domain, it will always be available. A third approach is to select technology that has achieved widespread use, which assumes that technology that is ubiquitous will remain available the longest due to its abundance.

Archivists may also preserve the hardware and software needed to reproduce their preserved content. This task presents the same difficulties as preserving media, because devices in storage will fail, even though they are not being used. Plastics used in the construction of hardware age and become fragile, lubricants dry out and become sticky, and electronic components called capacitors deteriorate with age, even when not in use. Digital devices suffer age related problems no differently than analog devices.

Some feel the complexity of modern digital systems presents such a challenge that sticking to simple analog systems is preferable. Simple systems undoubtedly offer advantages, but it's difficult to make such a broad generalization, and it's also not the case that all analog systems are simple. Others point to the incredible versatility of digital systems. This versatility allows content to be accessed on a broad range of devices including highly portable handheld devices. While the underlying technology may be complicated, the means of access is simple. It falls to the archivist, armed with knowledge of both types of systems, to study each case on an individual basis, and to make the best choices possible to preserve their information.

Jim Trafidlo
617-487-6188 Phone
jtrafidlo@nationalboston.com
115 Dummer Street, Brookline, MA 02446