Moving Beyond Manual Media Migration

James Lindner and Gilad L. Rosner
Media Matters, LLC
New York, New York, USA

Abstract

Magnetic tape is the world's most common media for audiovisual recording. Archives, broadcast networks, and consumers have used magnetic tape to store and play back their valuable audiovisual content because it has been inexpensive and plentiful. Some common professional magnetic media formats are BetacamSP, U-matic, DigiBeta, D1, D3, and BetacamSX; among consumers, VHS has been the dominant format. Cameras to record and players to play back magnetic media are also commonly available in both consumer and professional cost ranges.

Long considered advantages, the ubiquitous and inexpensive nature of the media also presents significant challenges to those concerned with the preservation of audiovisual heritage. Over the decades, more than 65 magnetic media formats have been introduced to the market. Many of these formats did not survive. With the demise of a format comes the obsolescence of its playback equipment, rendering countless hours of material unplayable.

In addition to issues related to the availability of magnetic media and its equipment, the physical medium itself has limitations. Though cheap and plentiful, magnetic tape is fragile and short-lived. Depending upon storage conditions, the life expectancy of modern magnetic tape is approximately 20 years. After that period of time magnetic media will degrade. From an archival perspective, twenty years is not a long time. The physical degradation of magnetic media poses considerable challenges to audiovisual archivists. Cheap and plentiful meets fragile and short-lived… so much media resides on cassettes and reels that no force on earth can save it all from degradation.

Enter migration: the only way to preserve content recorded on magnetic media is to migrate the information to new tape, or, thanks to recent advances in compression and declining storage prices, to digital media. Unlike simple video duplication, audiovisual migration is an esoteric and expensive art. Few companies can do it, and costs range from $250 to $500 per finished hour of media, with pricing based on the idea that all tapes are damaged when they arrive.

A review of traditional audiovisual migration is helpful here. Archives and other content holders first select a portion of their collection for migration. The tapes are inventoried, packed, and shipped to the migration company. Upon receipt, each tape is inspected and cleaned. It is presumed that all tapes will exhibit some form of deterioration and the cleaning process will reveal the extent of the damage. Tapes requiring little or no restoration continue through the migration process. The majority of tapes will only need minimal cleaning for playback. Tapes requiring extensive restoration are handled by specialists who use a variety of methods to resurrect the tape to a point where it can be played back and recorded onto a new master. In either case, humans perform all actions needed to migrate the content of a batch of tapes to new tape stock. A badly deteriorated tape could go through two to three specialists, each with their own skill level and area of expertise, before successful migration. Note that the end result of this labor-intensive process is one new master per source tape, and nothing more.

The approach to preserving tapes detailed above has three glaring problems. The first is cost. Specialized facilities, knowledgeable staff, and professional, properly serviced equipment are expensive to maintain. High costs are inevitably passed onto the customer. The second problem is scalability. A top-flight team of 2-3 specialists working 40-hour weeks could migrate approximately 5,000 hours of content per year. The processes involved when using human labor simply do not scale as you increase the size of the collection of tapes. Thirdly, the majority of tapes arriving at video migration facilities are in fair to good condition, contrasting with the established practice of treating every tape as if it were damaged. These three problems – coupled with the dangers of shipping a collection away from its home, the costs of insurance, the time to process a job, and the absence of meaningful metadata about the migration process – clearly illustrate where there is room for improvement in this process.

Current manual processes limit the preservation of our audiovisual heritage. Traditional methods simply cannot keep up with the vast amount of magnetic media stored in the world's archives. The costs are too high, the output too low. Without a fundamental change, much of the magnetic media holding the world's audiovisual heritage will disintegrate before its content can be saved.

Looking closely at the manual migration process, it is useful to analyze where human resources are spent. In a migration facility, human resources are spent on *every* tape. Rather than having trained engineers examine only the *problem* tapes, every tape is visually inspected. The pricing structure for a migration job reflects this inefficiency – tapes in good condition can cost as much to migrate as tapes that have extensive damage.

Lessons in Automation

The manufacturing industry model offers many lessons on increased productivity and the effective use of both financial and direct labor resources. If our aim is to maximize these resources and increase productivity, we should examine how automation technologies and equipment can improve migration processes. Pertinent automation advantages include economies of scale, modularity and reduced human labor.

Economies of Scale

The basic idea behind an economy of scale is the more product you move, the lower the cost to move it. Cost advantages are obtained through increasing the scale of a process, whatever that process may be. In the case of traditional audiovisual migration, processes do not scale because the labor required to inspect, clean, restore, and migrate each tape does not decrease with increased numbers of tapes. An increase in the number of tapes requires an increase in either the size of the workforce, or the length of time of the job. Attempts to increase the speed or quantity of tape migration without an increase of other resources results in a decrease of job quality.

Modularity

Technology changes rapidly. It's very important to be able to acclimate to change without having to re-educate your workforce, or make large capital investments in new equipment. Traditional audiovisual migration can handle changes in tape format, but little else. For example, adding the capacity to encode content while migrating it to new tape stock would require a substantial retooling of a migration company.

Reduced Human Labor

Human resources are amongst the most expensive components of any manufacturing process. In the right circumstances, reducing a company's human component can confer the benefits of streamlined processes, lower overhead, improved output quality, and greater consistency. In the case of traditional audiovisual migration, humans perform many repetitive tasks in the course of a 40-hour week., increasing the probability of mistakes occurring due to loss of focus or interest.

In the manufacturing industry, we see a culmination of the process advantages listed above in the concept of the *factory*. A factory may seek to combine and streamline processes, compartmentalize functions, create modular systems, increase scale, monitor quality control, reduce human labor and automate where possible. Through this it is hoped that they will create efficiencies, keep pace with changing technology, reduce cost and redundancy and achieve greater output. Fortunately, this approach can be adapted for the preservation of audiovisual archives.

The European commission PRESTO was launched in 2000 to address the long-term problems of audiovisual migration. They found that:

> …[t]he key to reducing costs without sacrificing quality is mass-transfers of material, using a 'preservation factory'. Software is needed to track all items moving through the factory to eliminate manual effort wherever possible. The factory approach, compared to conventional item-by-item preservation work, reduces costs by AT LEAST 50% -- and some PRESTO partners have demonstrated reductions of up to 75%.
>
> *European Project Presto Overview*

Removing Inefficiency and Improving Output

There are a number of inefficiencies in the traditional audiovisual migration business. It's useful to examine them in the context of the manufacturing discussion above.

Tape Handling

The production process is done exclusively by hand. Once cleaned, skilled operators insert tapes into playback decks, and destination tapes into record decks. The only piece of equipment used to affect the signal is a time base corrector (TBC), and human operators make judgments about signal quality and color accuracy using a vectorscope and waveform monitor. Problems of consistent tape handling, as well as staffing issues – absences, loss of focus, quality drops due to multitasking – contribute to inefficiency.

Tape Cleaning

Tape cleaning, is a laborious and delicate process. As magnetic media ages, the likeliest point of degradation is the adhesive agent holding the metal oxide to the plastic backing of the tape. In the case of advanced degradation, the cleaning process could actually exacerbate deterioration by creating more dropouts, or by seriously damaging the audiovisual signal.

Current cleaning machines are blunt and simple: they clean forward, and they clean backwards. In an attempt to reduce the potential for further damage, traditional migration companies pay a person to watch the cleaning tissue in the cleaning machine all day. When the tissue is clean, the tape is clean, but if the tissue never gets clean, it's a sign that the tape is "shedding," and further cleaning may damage the tape and compromise the migration of the information.; the danger of "over-cleaning" is very real. This means that a skilled person must observe the tissue constantly, and switch off the cleaner when she feels that the tape is in danger. This is a highly repetitive task, requiring constant attention and focus from a skilled technician whose skill set may not be fully utilized. A lapse of focus or attention is a real threat to the quality of the work. Lastly, customers often pay for cleaning on a per minute basis; any problem arising from this process, be it caused by advanced degradation or human error, will result in customers paying for more time on the cleaner than necessary.

Quality Control

This manual, real-time migration process results in the production of re-mastered tapes with potentially inconsistent quality, and little else. While some facilities maintain Manual Condition Reports, the value and consistency of these reports is subject to the focus and skill of the technician during playback. The reports may miss many crucial problems as many operators do not record their TBC settings on a per tape basis, nor do they make consistent notes about other adjustments made during the migration process. Opportunities for recording important data relating to the provenance of the media, and source tape health are lost if condition reports are not completed, or attention to detail is inconsistent.

Production limitations are inherent in this manual system. The most focused, attentive, skilled operators process a limited numbers of tapes during the course of their shift. In addition, periods of idle downtime engender a loss of focus by the highly skilled operators, further endangering job quality.

A Proposed Solution

One solution to the problems outlined above would be to introduce a factory approach to the migration of audiovisual material. By combining, streamlining and introducing automated processes and removing human resources where possible, quality would increase, costs would drop, and new services could easily be added.

The next step in automating the migration of magnetic tape would be to remove human resources from tape handling. Robotic technology that can reliably manipulate cassette tapes is, by now, an established, commodity. Used in broadcast environments, these robots perform repetitive tasks, such as alleviating the need for people to cue up and play commercials during television programs. Robotic tape handling is a straightforward technology that can be adapted to become the backbone of an audiovisual migration factory. Robotic systems configured to host professional video playback and recording decks are in use now. With the substitution of grippers and gears for human hands, migration processes can run twenty-four hours a day, seven days a week; systems can be mobile, or operate in fixed locations. By adding more robotic capacity, processes can scale without drastic cost increases. Furthermore, many robotic systems use interchangeable libraries to hold cassettes – this modularity ensures that robotic migration systems can accommodate the variety of formats found in the world's audiovisual archives.

An automated magnetic tape cleaner is the first step in streamlining the migration process. By replacing the tedious, repetitive tasks involved in the tape cleaning process with an automated electronic process, reliable results will be produced with consistent quality. By using sensors instead of human eyes, over-cleaning can be avoided, and defects in the tape can be identified and reported. Questionable quality can be reduced and customers of the migration service can be assured that they are not paying for unnecessary system time. The technology to accomplish all of this is available now, despite the fact that magnetic tape cleaning technology has not evolved much in 20 years.

Obviously, robots are only brainless tools without software. To create an automated audiovisual migration factory for magnetic media, customized software must be written to replicate decisions made by media migration technicians. Besides the complicated task of creating the instructions to carefully handle fragile audiovisual masters, logic must be designed to recognize problems in the migration cycle that could endanger the source tape, the destination tape, or the system itself. Expert knowledge must be synthesized and programmed into an expert decision-making system to create a truly automated system. For example, if a master tape sheds too much oxide during a cleaning process, control software should recognize the problem, stop cleaning, remove the tape, and put it on an exception report. If a recording is silent and shouldn't be, the control software should sense this, and attempt to re-record. Indeed, quality control relies on this type of dynamic, self-healing logic to ensure that tapes are migrated correctly the first time.

Another tantalizing benefit of advanced computer control and the logging of migration processes is the resultant metadata generated for each tape and for entire collections. Where manual processes produce inconsistent metadata, automated systems can be programmed to produce a wealth of consistent information. The diversity and amount of metadata enables the evaluation of a collection's health in ways never before available. For example, archivists can peer into the health of tape batches segmented by age, source, tape length, and storage location. This data can be culled and exported in any format, including XML, a powerful and flexible language for cataloguing, research, and publishing to the web.

Triage

While migration can be automated, not all tapes are suited for automatic migration. Tapes must first be visually inspected to ensure their survival through the migration process, and to ensure that they will not cause damage to the migration system itself. Some tapes are too fragile or damaged to go through an unattended process; those tapes are, thankfully, in the minority. Some tapes will pass a visual inspection, but could still cause problems during migration. An intelligent system will weed these tapes out by detecting serious problems during the cleaning or migration cycle. Anecdotally, it is expected that 5% of tapes will not be candidates for automated processing, and will need to be handled by experts using traditional migration techniques. The remaining 95%, however, can benefit from less expensive automated migration. The concept of "triage" is most applicable here: evaluate which tapes can be saved easily, and send the rest to specialists.

Crossing the Digital Divide

An automated audiovisual migration system should look forward to encompass technological advances. While most archives' infrastructures have been based on magnetic media, the push towards digital archives is inevitable. Hard drive and processing costs continue to decline, and advances in encoding software and hardware maximize hard drive space while improving image quality. Consequently, a modern migration factory must incorporate encoding technology into its mix of services. While the focus of modern automated migration practices could be focused solely upon preservation, it can also be a tool for improved access to audiovisual collections. PrestoSpace, the commission that succeeded PRESTO to implement its recommendations, states:

> The major challenge that all audiovisual archives are facing now and for the next decade is the preservation of their legacy materials and the migration to the digital environment… The related challenge is to provide wider access to the contents through adequate technology, services and rights management. The whole [PRESTOspace] project is focused on this challenge, because access – specifically, radical improvement in access – is seen as the only substantial motivation for preservation funding. The basic idea of a *preservation factory* is that it **delivers access**, by integrating within the preservation work ALL the functions and processes to exploit the new forms of access (principally web access) available to digital material.
>
> *PrestoSpace Manual [original emphasis]*

Inevitably, magnetic tape technology will become less attractive to maintain. Video recording and playback decks are high-priced when new, and expensive to maintain when old. As tape formats die, locating and maintaining legacy decks is a difficult and costly proposition. Finding experienced engineers to repair and maintain equipment becomes harder with each passing year. With declining costs, the improving reliability of hard disk-based storage, and technologies such as automated migration factories, the pressure to supplant tape-based archives with digital ones will continue to increase.

Just as with tape formats, encoding technology must be modular to keep pace with changing codecs and file formats. The variety of uses for digital files – web delivery, DVD delivery, preservation copies, etc. – calls for an inherent flexibility in the migration factory's system architecture.

Observations and Conclusions

The authors of this paper have spent several years prototyping an automated migration factory known as SAMMA, the System for the Automated Migration of Media Archives, and have learned a great deal along the way. As conventional archival migration processes have, for the most part, been fulfilled by off-site vendors, most archives have not had technical, hands-on experience preserving their media. Audiovisual archives poised on the edge of the Digital Divide are faced with significant changes in the preservation paradigm, changes requiring a diversification of the traditional role of the archivist. Archives' knowledge of technology as it relates to digital migration and related preservation issues is limited, but the need to address technological innovation is monumental. Archives must acquire an understanding of this technology, and be shown how using automated, or semi-automated systems would affect their well-established workflow in a positive way.

From the authors' observations, a lack of understanding of digital technology – in particular, of encoding and storage – is a common stumbling block for nearly all archives. Even the use of simpler digital technology like databases for cataloguing collections is far from universally implemented or understood. Moreover, it should be expected that few archives are prepared to store and manage large numbers of digital files. The infrastructure simply isn't there, and the costs and resources of creating one is a significant challenge.

Transformation to digital preservation presents additional challenges in the area of copyright and rights management Most archives do not hold the copyright for the programs in their collections, and are (appropriately) concerned about access issues related to digital files. A poor understanding of digital technologies combined with an even poorer understanding of Digital Rights Management, or DRM, leads archives to remain conservative about digital technology.

We recommend detailed discussions occur early in the relationship with an archive to ensure both sides have a clear understanding of the recommended technology and project objectives. A granular explanation of the nature of digital technology, present and future costs, and organizational impact must occur at the early stages of automated migration systems deployment. The degree to which technology is understood and adopted varies with each archive. Issues such as pre-existing barcodes, databases, digital storage, quality control, and workflow must be discussed during the ramp-up to deployment. No automated migration system can be "off the shelf" – the particulars of preservation, restoration, storage and archive management, content description, delivery and access demand a tight integration with a migration system.

Automated migration systems are being created now. The road to broad adoption of these new technologies has numerous obstacles, not the least of which are limited customer understanding, institutional inertia, and on-going technology maturation. However, the world's audiovisual heritage is endangered, and innovative preservation methods are called for to address this impending disaster.