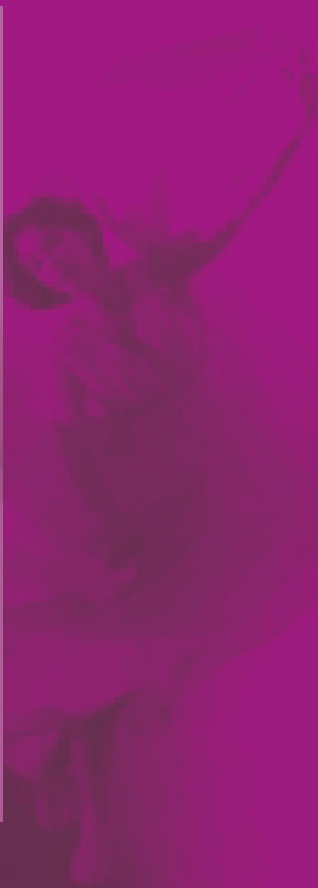


DIGITAL VIDEO PRESERVATION REFORMATTING PROJECT



DANCE HERITAGE COALITION, INC.

DIGITAL VIDEO PRESERVATION REFORMATTING PROJECT

A REPORT

Prepared by Media Matters, LLC for the Dance Heritage Coalition

Presented to The Andrew W. Mellon Foundation

June 2004



The Dance Heritage Coalition (DHC) was founded in 1992 to address the problems that were identified by a study of the state of preservation and documentation of dance in America. Jointly commissioned by The Andrew W. Mellon Foundation and the National Endowment for the Arts, the resulting study *Images of American Dance* recommended the formation of an alliance of the nation's major dance collections (1) to facilitate communication; (2) to develop national standards, policies, and priorities; and (3) to implement collaborative activities and projects in the fields of dance preservation, documentation, and access. The DHC's mission is to make accessible, enhance, augment, and preserve the materials that document the artistic accomplishments in dance of the past, present, and future. It also now serves as a think tank and convener for the dance heritage field.

Member Organizations of the Dance Heritage Coalition

American Dance Festival

Dance Notation Bureau, Inc.

Harvard Theatre Collection, Houghton Library, Harvard University

Jacob's Pillow Dance Festival

Library of Congress

New York Public Library for the Performing Arts, Dance Division

Lawrence and Lee Theatre Research Institute, Ohio State University

San Francisco Performing Arts Library and Museum

*Anacostia Museum and Center for African American History and Culture,
Smithsonian Institution*

© 2004 Dance Heritage Coalition and Media Matters, LLC

This report was written by Media Matters, LLC, a technical consultancy specializing in archival audio and video material. Media Matters provides advice, analysis, and products to media archives that apply the beneficial advances of technology to collection management. James Lindner, Managing Member of Media Matters, acted as Principal Investigator for this project, assisted by Justin Dávila, Aron Roberts, Gilad Rosner, and Jennifer Crowe. Support for this project comes from The Andrew W. Mellon Foundation and the National Endowment for the Arts.

DANCE HERITAGE COALITION
1725 Eye Street, NW Suite 300
Washington, D.C. 20006
(202) 530-3209 / (202) 530-3211
info@danceheritage.org
www.danceheritage.org

MEDIA MATTERS, LLC
500 West 37 Street, 1st Floor
New York, NY 10018
(212) 268-5528
info@media-matters.net
www.media-matters.net

CONTENTS

Preface	5
Introduction	7
Why Study Dance?	9
The Current State of Dance Video in America's Archives and Libraries	10
The Digital Video Preservation Reformatting Project	13
Defining Preservation Quality for Dance Archives	16
Traditional Methods for the Preservation of Video	26
Innovative Ideas for the Preservation of Video	30
The Determination and Specifications of Preservation File Format Candidates	33
Lossless Compression	34
Lossy Compression	37
File Wrappers	39
AAF	39
MXF	42
MXF vs. AAF	43
Construct the Software (If Necessary) to Create Preservation File Format Candidates	45
Produce a Footage Test to Include Dance Footage and Other Test Footage	46
Methodology	47
Compression	51
Codec Analysis	54
MJPEG2k	54
MPEG-2	54
MPEG-4	55
Windows Media	55
RealMedia	56
QuickTime/Sorenson 3	56
The Analysis of the Tests Run on the Footage	57
Summary Analysis and Recommendations	99

Appendix: Analytic Tool—Genista’s Media Optimacy	104
Video Quality Metrics	104
Relative and Absolute Metrics	105
Metric Type Description	107
Perceptual Metrics	107
Jerkiness	107
Blockiness	108
Blur	108
Noise	108
Ringing	108
Colorfulness	109
Watermarking Artifacts	109
MOS Prediction	110

PREFACE

During the winter of 1999 and through the spring of 2000, the Dance Heritage Coalition (DHC) sponsored a series of meetings known as the National Dance Heritage Leadership Forum. At these gatherings, dozens of professionals from both inside and outside the field of dance heritage articulated mandates for advancing dance documentation and preservation during the next ten years. Included was the plea that the DHC launch a national campaign to address the magnetic media crisis—a crisis that has already meant the loss, through deteriorating videotapes and format obsolescence, of many of the moving images that are the record of this nation's diverse, dynamic history of dance.

In response to this directive, the DHC called a meeting in July 2000, moderated by Carl Fleischhauer of the Library of Congress, to lay out a plan for a project to migrate analog videotape to digital for preservation purposes. In the spring of 2003, the DHC was awarded a grant from The Andrew W. Mellon Foundation to examine the technology, which would lead to establishing standards for the preservation community. Our work was completed in the spring of 2004, with the recommendation to use JPEG2000 and Material Exchange Format (MXF) as the file standard. The dance community has every reason to be proud. Much to the surprise of many in the archival community, the field of dance initiated this work. The results will impact areas far beyond the performing arts. (In July 2004, Digital Cinema Initiatives, a joint venture of Disney, Fox, MGM, Paramount, Sony Picture Entertainment, Universal, and Warner Bros. Studios announced that they had also chosen JPEG2000 as their standard.)

The story does not, of course, end here. Funding must be secured so that the larger repositories may begin the work of reformatting their holdings; funding is also necessary to maintain digital files. Hubs need to be established so that independent choreographers and dancers as well as smaller

organizations can avail themselves of this technology. Clearly, there is still much to do. On behalf of the DHC, I can promise this will be a priority for the future—a more secure future for the thousands upon thousands of videotapes that document our dance heritage.

ACKNOWLEDGMENTS

On behalf of the DHC, I wish to extend warm thanks to Carl Fleischhauer of the Library of Congress, who, as Principal Advisor, offered the original stimulus and advice for this project. The National Endowment for the Arts provided funds for the first meeting, *Designing an Experiment in Digital Video Reformatting*, held in July 2002 and the DHC recognizes with gratitude the Endowment's continued support of documentation and preservation projects. The Dance Division of the New York Public Library for the Performing Arts, Madeleine Nichols, Curator, and the Division's moving image specialist Else Peck spent hours assisting in the selection of video clips as did Norton Owen, Director of Preservation at Jacob's Pillow Dance Festival. As Principal Investigator for the project, the DHC is, indeed, fortunate to have engaged James Lindner of Media Matters, LLC. A renowned leader in the field of moving image preservation, Mr. Lindner and his colleagues Justin Dávila, Jennifer Crowe, Aron Roberts, and Gilad Rosner at Media Matters, LLC patiently explained technical issues and gracefully accepted my slow, but gradual understanding of the world of digital compression. Finally, the DHC is profoundly grateful to Donald J. Waters, Program Officer, and Suzanne Lodato, Associate Program Officer, Scholarly Communications at The Andrew W. Mellon Foundation for support of this project.

—*Elizabeth Aldrich*, Executive Director
Dance Heritage Coalition

INTRODUCTION

During the 1990s, many organizations began the digital reformatting of their library and archive collections. *Digital reformatting* refers, broadly in this context, to the work carried out by various types of projects. At one end of the spectrum were projects with the principal goal of increasing access to collections; in many of those cases, the making of preservation copies was a secondary goal or even an unacknowledged outcome. At the other end of the spectrum were projects intended from the start to make preservation copies, understood to be copies that served the same functions that were previously performed by microfilm (for printed matter or manuscripts), by copies on continuous-tone film (for prints and photographs), or by copies on magnetic tape (for sound and video collections). Roughly speaking, preservation copies were and are intended “to take the place” of the originals if the need arises.

The barriers in the use of digital technology to reformat library and archive content have fallen. Not surprisingly, relatively simple entities like the printed pages of brittle books were the first to be explored. Soon after came the creation of surrogate images for pictorial materials. As the technology became available to the library, archive, and museum world, reproduction quality increased markedly. By 2004, the digital copies surpass their analog-film predecessors in terms of reproduction quality. The development of better online delivery technologies broke the barrier for maps, and now many libraries are reformatting large color sheets, foregoing the one-map microfiches that were formerly created. The most recent barrier to fall has been in the area of sound recording; it is now easier to make digital-file copies of sound at very high resolution, and it is increasingly practical to sustain large audio files in server-based storage systems.

This report focuses on the next barrier we face: video recordings. It highlights a variety of challenges that remain, explaining nuances and intricacies in language that is informative without being so technical as to be ob-

scure to nonspecialists. The story told here demonstrates that the digital reformatting of video recordings is both a science and an art, in a state of becoming. We owe the Dance Heritage Coalition a grateful nod for organizing this effort and for sharing its findings with colleagues worldwide. It is exhilarating to read this opening act in our video reformatting drama, even as we recognize that several more acts must follow before the drama is complete.

—*Carl Fleischhauer*, Project Coordinator
Office of Strategic Initiatives
Library of Congress
Washington, D.C.

WHY STUDY DANCE?

In centuries past, and continuing into the present era, there has been a tremendous flowering of creativity in all areas of dance, including ballet, modern dance, social dance, Native American dance, folk dance, tap dancing, and dances linked to jazz. Comprising an entire world of spiritual and secular ideas, stories, emotions, and human experience, dance (and its accompanying music) is part of our shared cultural experience and heritage. We document dance so that everyone can explore it and thereby better understand its meaning.

Dance itself, however, is intangible. Only its artifacts, such as programs, photographs, costumes, and set designs live on in a tangible form. While still photographs can capture some aspects of performance, dance movement could only be captured when the technology to record it became available. Many of the earliest motion picture films featured extensive dance scenes, such as D.W. Griffith's silent classic *Orphans of the Storm* (1921). With such filming, dance was an art form that could be saved as well as shown to large audiences.

Since the introduction of videotape technology in the late 1950s, dancers, choreographers, dance companies, and those capturing dance as part of anthropological fieldwork have increasingly relied on videotape to record and replay this ephemeral art form. When videotape recording was first introduced, successful operation of the technology was beyond most. In addition, access to this equipment was very limited. In the mid-1960s, however, videotape equipment became more compact, less expensive, and easy to operate, allowing broad application. Thus, it became possible to use video to capture live performance. From that time video technology has played important roles in the dance community; it enables dance to be recorded for a variety of purposes—for documentation, for the creation of choreography, and for various performances purposes.

THE CURRENT STATE OF DANCE VIDEO IN AMERICA'S ARCHIVES AND LIBRARIES

Magnetic tape has provided a medium to record and replay dance history at will, and it remains the most common method of documenting all forms of dance. Only recently has the dance community realized that, in fact, analog videotape is as ephemeral as dance itself.

In 2003, the Dance Heritage Coalition (DHC) created the National Dance Heritage Videotape Registry, a database containing detailed information on the videotape collections of dancers, choreographers, dance companies, dance teachers, museums, dance festivals, presenting organizations and performing arts centers, management organizations, libraries, colleges and universities, videographers, and producers.

The Registry suggests that the 300 respondents to a detailed questionnaire (distributed by the Dance Heritage Coalition) hold more than 180,000 videotapes, recorded between 1956 and 2003. This sampling is but a minute representation of the entire field in North America and worldwide; there are literally hundreds of thousands more tapes, many of which are endangered by a number of factors, including format obsolescence (whereby the playback equipment is no longer readily available), as well as the chemical and physical deterioration of the actual tapes.

The results of the National Dance Heritage Videotape Registry questionnaire indicate a burgeoning magnetic media crisis. Urgent steps must be taken. More than 25% of the respondents believed that at least some of their tapes were physically damaged. More than 50% did not have the information and/or the staff to evaluate their collections. More than 80% have no procedures in place at all to ensure long-term preservation of their tapes. The number of aging tapes in dance archives will only increase with time. There were 11% of survey respondents with videotapes that were

recorded between 1956 and 1970; 55% have videotapes recorded between 1970 and 1985. More than 50% of respondents lack playback equipment for all the various tape formats contained in their collections. To compound the situation, large institutions with large budgets, such as the New York Public Library for the Performing Arts and the Library of Congress, have expressed concern regarding the longevity of playback machines. Meanwhile, the small dance archives are in much the same situation, and they have very few resources to maintain their few playback machines.

Preservation experts strongly encourage the migration (re-recording and re-formatting) of endangered analog videotapes to a format such as Betacam SP. However, the cost of Betacam SP is as yet too prohibitive for most dancers, choreographers, and dance companies. To help in this situation, during the winter of 2004, the DHC provided funds to reformat approximately 70 at-risk videotapes to Betacam SP. These included the work of American dance icons Ted Shawn, José Limón, Lew Christensen, Harold Nicholas, and Gregory Hines, to name a few. Regrettably, no playback machinery could be found to reformat Meredith Monk's original cast performance of her seminal work, *Education of A Girlchild*, recorded in 1973, or the 1976 videotapes of Anna Sokolow's *Deserts* and her *Lyric Suite*. The only record of modern dance pioneer Lester Horton's technique, as demonstrated by Horton dancer, Bella Lewitzky, has completely deteriorated and cannot be migrated. These performances—important milestones in the legacy of American modern dance—are now lost forever. Without a concerted preservation effort, the dance world is in danger of losing many more of the moving images that have become the iconic and collective memory of all forms of twentieth-century dance.

The problem, however, is not only the old analog recordings. Many of the tapes being recorded today are “born digital,” meaning that the technology used to record them is digitally based. While such digital recordings have advantages, they also have very significant preservation challenges (especially those concerning compression). When they are added to an already complex matrix of preservation challenges, the result may overwhelm our

current capability to ensure that our dance heritage survives. The risk, then, is not only to our legacy analog recordings but also to our modern digitally born recordings.

THE DIGITAL VIDEO PRESERVATION REFORMATTING PROJECT

The Dance Heritage Coalition has closely monitored the impact of the development of digital technology on the dance community, beginning in the mid-1990s. In a report to the National Endowment for the Humanities in 1997, the DHC identified a critical need for the preservation of moving image and audio materials, particularly for dance recorded on videotape.¹ Digital preservation of these materials was and continues to be an area of interest for the DHC. A Technical Advisory Group was created in 1998 to guide and inform the DHC in these matters, and thus the preliminary structure for the Digital Video Preservation Reformatting Project was born. Drawing upon professional expertise in moving-image video migration, the group proposed using the dance community's difficulties with video preservation as a model to address the complex issues surrounding the preservation of magnetic media as a whole.²

¹ The members of the Dance Heritage Coalition participate in various organizations that are leading the way—nationally and internationally—in providing guidance and standards for preserving, documenting, and accessing America's cultural heritage through digital means. The Coalition is able to shape its initiatives and develop strategic policies, in part, through its members' involvement in this vanguard of technology organizations and working groups. These include the Digital Library Federation (DLF), Research Library Group (RLG), the Coalition for Networked Information (CNI), and Internet2. The DHC frequently consults with organizations such as Association of Moving Image Archivists, Bay Area Video Coalition (BAVC), Heritage Preservation, Image Permanence Institute, as well as leading video preservation experts Sarah Stauderman (Smithsonian Institution), James Lindner, and William T. Murphy (formerly of the National Archives and Records Administration.)

² Members of this Advisory Group have included Wes Boomgaarden, Director of Preservation, Ohio State University; Carl Fleischhauer, then with the National Digital Library, Library of Congress; Gerry Gibson, then with the Library of Congress; Steve Hensen, Special Collections Library, Duke University; Catherine Johnson, former director of the Coalition; Madeleine Nichols, Curator, Dance Collection, the New York Public Library for the Performing Arts; Vicky Risner, Head of Acquisitions and Processing, Music Division, Library of Congress; Abby Smith, Director of Programs, Council on Library and Information Resources; and Jim Wheeler, Belmont, California.

The Dance Heritage Coalition has been well aware that it is not just the dance community that is troubled by rapidly deteriorating videotapes. During the discovery portion of the project (Phase I), the DHC found that in the commercial, academic, and public spheres the body of data required to make informed decisions about *how* to proceed with an effective digitization program was surprisingly scattered. Many diverse communities were examining bits and pieces of the video preservation puzzle, but few solutions showed promise specifically for the dance field. With funds from the National Endowment for the Arts, the DHC called a meeting in July 2002 to discuss the possibility of designing an experiment to explore the most appropriate method of transferring analog videotapes to digital files for preservation purposes. To do this, a variety of dance videotapes would be used in the tests.

The result of the July 2002 meetings was the *Digital Video Reformatting Preservation Project, Phase I and II*. (Phase I, the discovery phase, is described above.) The report of those meetings suggested several directions for exploration.³ Phase II was defined to examine the suitability of a variety of popular digital-compression types as a potential preservation format, by applying them to various types of dance footage found in dance archives. Phase II also examined the behavior of these new files within so-called file wrappers, a technique used to hold both essence information (picture and sound) with metadata (information about information—in this case condition or other descriptive information). It is desirable, as expressed in the Dance Heritage Coalition's Winter 2003 project proposal to The Andrew W. Mellon Foundation, that "the digitization process will not only conserve the original object, but will reduce the further deterioration of (and provide access to) rare, fragile, and vulnerable materials. By setting preservation standards, the outcomes expected from this project will have enormous resonance not only for the dance community, but also for every major archival institution."

³ The report is available from the Dance Heritage Coalition.

The findings of Phase II are presented here in this report. They include technical experiments on an assortment of dance footage, to determine the merits of a variety of compression and storage schemes for the preservation of analog video dance footage as digital files. In addition, this report suggests a potential preservation strategy for the dance community, based on a consideration of the test results, the analysis of industry trends that have been in place for some time, and the new possibilities presented by recent trends in both standards and hardware.

DEFINING PRESERVATION QUALITY FOR DANCE ARCHIVES

The July 2002 committee identified the following three categories of pass–fail factors for preservation copies. The test will apply these factors to the degree that is practical.

The *quality* of the picture and sound, including resolution, chroma bandwidth, luminance, synchronization pulse, and a lack of phase shifts. A copy will pass the quality test if the measurement of these elements shows little or no diminishment or degradation when compared to the measurements of the original.

The *usability* of the end product or the resulting preservation master copy or the working copies made from that master must support the following performance measures:

- a. It must be possible to edit the copy.
- b. The copy must retain any information that allows users to run processes on the footage, such as search engines.
- c. The copy must allow output that can produce an HDTV (high definition television) copy.
- d. The copy must permit tape-to-film transfer, and it must allow freeze framing. (Freeze-frame capability is important for the dance community, since users must be able to view single frames clearly, to study details of choreography.)

Preservability of the end product (i.e., end product must be migratable and must avoid technical protection, such as encryption). The format must also be open source, public, well documented, and should carry no fee or very low fees.

In short, the idea of the committee was to define a level of preservation quality that captures the essence (picture and sound) of dance recordings in such a way that the copy is essentially unchanged from the original, if possible; or if that was not possible, to have the change be extremely minimal. The most important concept was that “a copy will pass the quality test if the measurement of these elements shows little or no diminishment or degradation when compared to the measurements of the original.”

This quality test is an extremely difficult technical challenge from a number of perspectives. Perhaps the most important is that for a high-quality copy to be possible, one would assume such a process to be already common in the broadcasting industry. This, unfortunately, is not true and never has been. For this reason, it is important to explore the notion of video quality, as well as to investigate the different technologies used to compress and distribute video.

Historically, providers of broadcast television and digital video content have been primarily interested in the way a picture looks when it is delivered, at the time of transmission or playback at the receiver, which may be a conventional television set or a computer monitor or other technology receiver. Images are delivered to different audiences in various ways. A few of the “traditional” techniques that have been used include transmission over the air as a terrestrial broadcast, by cable TV, or via satellite. More recently, images and sound have been sent electronically, as data, which then can be sent as files to a remote location, to be played there or transmitted as a continual data stream over the Internet or for a computer screen at a kiosk.

In general, the goal is to deliver video of viewable, useful quality. Note that we did not say that the goal is to deliver “ultimate quality” or “superb quality” but useful quality—and, in particular, useful quality for the intended purpose or application. In fact, there is not yet a single picture-quality level, and this has always been so, throughout industrial broadcast history. When defining preservation quality, one must be aware of the tremendous

diversity of picture quality in the first place. Since there is, as yet, no single quality level for which to aim, any preservation strategy must account for that tremendous diversity, both in the form of the image and its intended avenue of distribution. Although there are standards to which a signal must conform, for proper viewing reception and reconstitution, this has little to do with the actual or perceived image quality. For example, an image of acceptable quality on a small window or computer screen, when the signal is being streamed and may be losing frames, will be of totally unacceptable quality when viewed on a high definition projected television screen in a theater. Thus, the expectations of quality must be scaled to the original, and to be efficient, any approach for preservation must be similarly scalable.

From the beginning of broadcast television (and even earlier during the decades of its development), many techniques have been used to try to balance the quality of an image delivered versus the cost of delivering that image.

When defining preservation quality for dance, we must be mindful of the larger technological world in which we live. That is to say: the technology used to capture dance is not unique technology; it shares the same heritage and equipment that is used for other applications, both industrial and private. Since the dance community must use the available technology when seeking to define preservation quality, we must keep in mind the constraints of the broader technological landscape. We must first carefully explore the technologies already used for image storage and distribution, because they will have to be used by the dance community and by others as well. It is unlikely that a “special” technology will be developed for the dance community, and even if possible, being on a technology “island,” isolated from the rest of the world, is of questionable value from a preservation point of view. To have important content on “orphan” formats or technologies has already shown to be a strategy of little value.

Preservation needs have never been issues embraced by electronics manufacturers—and this makes the current challenge all the more difficult. Manufacturers make money by selling new equipment, not by making equipment (with the replacement parts and accessories) that will last for centuries (even if they could). Therefore, when discussing the preserving of image quality for dance, we must explore and consider the broader technological landscape, with the tools that are now used. For this reason, a key element of Phase II was to examine the technology, specifically the video compression technology.

Video compression is, in fact, a series of techniques used in recording or playing back video imagery that conserves valuable, often expensive resources. For example, the resource that is most frequently saved is storage space; a file that is compressed takes up less space on a computer hard drive than a file that is not compressed. Video compression techniques can be used to conserve other resources, which include (1) bandwidth (one can think of that as the capacity of a computer connection to carry information); (2) time (the time it might take to download or copy a file), or (3) cost (smaller files use less hard drive space or other storage, which costs money—so less space often means less money).

In the context of defining “preservation quality,” video compression must be viewed as a process of compromise. The process of video compression comes at a price. Sometimes that price is the literal cost of the hardware or software that provides the compression (which is called a *codec* or *coder/decoder*). At other times the cost is for the computer power that is required to make the compressed file, or in the time it takes to make such files. The biggest compromise, however, is often taken in image quality. Because our eyes are not sensitive to detail when objects move on the screen, (the brain assumes, or fills in, the expected details), video compression techniques frequently use shortcuts in image quality for the purpose of saving space. Redundancies—for example, a detail that is repeated—are frequently removed; removal allows space to be saved. There are other tradeoffs

(discussed at length below), yet the important concept is that video compression is a series of techniques that allow for savings—but also come at a serious cost. The cost frequently is in image quality.

Broadcasters and online providers have become experts at tweaking digital video compression algorithms in order to deliver previously enormous files as smaller files. They accomplish this by creating parameters for acceptable levels of video signal loss, eliminating just enough video information to fool the human eye and brain into thinking that what it is seeing on the screen is a decent, coherent, and consistent picture.

Archives, and dance video archives in particular, may not have this luxury. Both archives and broadcasters are interested in providing access to video via low-bandwidth digital files, but for archives the institutional mandate is one of preservation, not merely content distribution. For dance archives, the stakes are even higher, since the analog footage in dance video archives is primary material, the history of the field. Analog footage provides a rich visual record of the output of the field of dance, and the taping has flourished without the benefit of large commercial, or even large non-profit, budgets. The dance community has thus created thousands of tapes, and it managed to keep up with the ever-changing formats and equipment.

The Committee has defined three factors for the investigation of digital video encoding schemes: image quality, usability, and preservability. The overall goals and desires expressed by the Committee were (1) to limit compression artifacts and obtain the best quality of image possible, while (2) expanding access to end-users and extending the portability of the file itself, within current and future archival systems.

Image quality means how good the recorded image looks to the human eye—and also to objective computer analysis. A digital video file format will pass the image quality test if post-compression measurements are a match, as closely as possible, to the original or reference source material. Ideally, they would be identical. If the digital, compressed file matches the

original file in a variety of areas—luminance, chrominance, synchronization pulse, lack of phase shifts, and others—with little to no degradation, it will be considered a successful candidate for preservation. This is not as simple as it sounds, as our results showed. Some techniques do a better job than others, depending on the source material and the quality that, in fact, varies from frame to frame in most video compression techniques. (This is discussed later in the report.)

The goal of any preservation effort can be thought of, ultimately, as to “do no harm” to the source materials you are preserving, and, in the specific context of dance recorded as video imagery, to have the copy not be “harmed” or different from the original. Archives should be able to use this footage in their current systems and the footage should be of high enough quality, with as much information as possible remaining intact, so that it may be used in future systems. To this end, it is desirable to create a preservation protocol that maintains the *usability* and the inherent value of source materials for future historical analysis. A preservation file format should maintain the highest level of usability possible.

Usability also refers to the way that information about the contents of a videotape can be described, so that it can be found by catalogs and by online search engines. The value of an archive is directly linked to how information therein is described. If information describing an archival object cannot be accessed, its value within the archive is diminished.

Currently, someone can type “George Balanchine” into a search engine on the Internet or a library catalog computer and get back a list of dances by George Balanchine, texts by George Balanchine, publications focusing on him as a subject, and anything and everything that contains the text metadata words “George Balanchine.” In the future, new technology—akin to facial recognition software—may be integrated into a search engine. If you feed the search engine a picture of George Balanchine, not only would it give you every Balanchine dance, but every video in the collection in which he appears (individual dances, symposia, other kinds of perform-

ances), whether or not he appears in the textual metadata. This could be an invaluable tool to researchers interested in painting a larger picture of a choreographer's life, for example. In order to take advantage of emerging search technologies based on image identification and to allow for ever more advanced technologies that will process dance and related imagery, *the highest level of video quality must be maintained when digitizing*. If detail in the footage is lost in the digitization process, it renders these technologies potentially useless.

The ideal file format candidate for the preservation of dance footage must not only maintain high levels of image quality and usability but must also enable the contents to be preserved over the long term—it must have a high level of *preservability*. Technology is constantly developing. Formats become obsolete, computer platforms come and go, and new methods are devised; therefore we must strive to find a file format that is flexible enough to survive for decades.

The chosen format should be nonproprietary—that is, not owned by an individual or a single company. Rather, the file type should have wide industry support and must allow for easy exchange between a wide variety of proprietary and nonproprietary types of systems. Users will need to perform a variety of operations with the files: editing on one system, adding graphic elements on another, creating special effects on another, and so forth. At present, it can be very difficult to convert one vendor's file type to another; therefore, there is a high level of interest in a file type that can interoperate among a variety of vendors' systems. Ideally, end-users should not need to purchase a license to employ the format.

When discussing preservability, we are also referring to any chosen video compression scheme's ability to pass the quality test at a level higher than that of visually perceived quality. While the perceived level of visual quality is extremely important, it is not the entire story. It is entirely possible in some situations, in fact, to fool the eye so effectively that while the images may look identical, the data representing them are, in fact, largely differ-

ent. As such, that data would fail our preservability test: an image may look good but it is not an accurate representation of the original data. Thus, it will have failed the preservability test. One may reasonably ask “why is this test important?” The reason is the test of whether it “looks good enough” might fail other levels of quality needed for additional types of analysis in the future, or it may fail a test of authenticity or artistic intent. For example, a codec may reduce background visual “noise,” which may actually be a visual distraction in many types of video imagery. This same background noise, which some may be able to distinguish and others may not, can in fact be part of the visual texture of a piece and the artistic intent of the author. Therefore, the act of changing that aspect—while, perhaps, being visually identical to some—has failed the preservability test.

Video footage, especially dance footage, presents many challenges to archivists. An example is the prevalence of both consumer and so-called prosumer-grade video recordings in dance archives. Formats such as VHS and Hi-8 are ideal for recording and playing back video signals for some archives. Compared to film, these formats simplify the necessary job of documenting the output of dance companies, festivals, and other events, while keeping budgets under control. By using these formats, a dance archive of modest means can easily amass a large collection of one-of-a-kind recordings, invaluable to dance scholars and aficionados. VHS and Hi-8 tapes (the former introduced in 1976, the latter in 1989) and camera equipment were inexpensive and, in their heyday, were easy to work with and plentiful.

Unfortunately, the signals recorded on VHS and Hi-8 tapes are inherently unstable, from a technical point of view, as compared with more expensive professional formats. In order to utilize these consumer and prosumer-grade materials in contemporary editing systems, it is first necessary to convert to a higher playback standard, to repair any signal instability. Also, to edit these tapes to any format other than VHS, for example, a conversion must also be made. Conversion does not inherently change how the signal looks, since VHS footage will still look like VHS footage, but it brings the

signal into compliance with the RS170A, or professional broadcast standard, so that it can be viewed and edited on broadcast-quality equipment. For the purposes of this study, we began our technical analysis of all videotaped materials by first converting tapes to RS170 broadcast standard. Such a conversion allowed the footage to be edited, as well as to be freeze-framed cleanly on a monitor for detailed scholarly analysis—of particular interest to the dance community. Without clean frames, analysis of the slightest movement, from the delicate hand gestures of Balinese dancers to the colorful waves of a Flamenco dancer's skirt, would be difficult to achieve with accuracy.

Since the 1980s, digital technologies have been developed at an exceedingly rapid pace in almost every area of communication, education, and recording. The basic technology behind broadcast television, however, has changed very little since the 1940s. The Federal Communications Commission (FCC) drew up a plan in 1997 that mandated broadcast stations to broadcast digital-only signals by 2006. So far, the PBS, Fox, CBS, ABC, and NBC networks have all adopted these standards, and they broadcast digitally in all major markets. Digital television will change the way we look at and listen to television. Not only will it expand the type of content that can be disseminated along with video, it will free up parts of the electromagnetic spectrum for other uses. The most obvious advantage of high-definition broadcast TV (HDTV) is the dramatically increased quality of the image seen on the screen. HDTV has up to six times the resolution compared to a standard (NTSC) signal. The images are very crisp, the detail is very fine, and perception of three-dimensional depth very pronounced, when compared to traditional standard-definition television.

High quality, detail-rich images will thus become ever more valuable in the world of digital television. The ante has been raised, and broadcasters are responding to the challenge accordingly. *When*, not *if*, archives rich in historical analog video migrate their collections to digital for preservation purposes, fitting these materials into the larger context of a high-definition broadcast world must be planned for in the overall strategy. For this reason,

it makes little sense to use compression schemes that seriously damage image detail when digitizing archival video footage. Such schemes essentially cannibalize the originals and lessen the value of the footage, in order to allow it to fit into storage solutions that, in time, will inevitably become less and less expensive.

For the purposes of this study, then, the ideal preservation format for dance footage must take into account the imminent demand for high-quality images. When archival dance footage is ultimately digitized, it must be done at the highest quality possible.

TRADITIONAL METHODS FOR THE PRESERVATION OF VIDEO

Since the 1970s, audiovisual preservation has advanced in small increments. The reliance has been on established technologies and methods to stem the tide of magnetic media degradation.

One option, in dealing with the overwhelming amount of audiovisual material, has been simply to do nothing other than control the environment in an effort to slow deterioration. This approach to preservation prescribes that all tapes be carefully climate-controlled, to slow as much as possible the degradation of the collection. Old tape decks and playback equipment would be stored, while archivists literally pray that replacement parts and skilled technicians will be available in the future. In this manner, waiting and seeing and hoping for the best, an archive might struggle until the inevitable death of its tape collection.

Approaching preservation in this manner, or “hoping for the best,” never really deals with the volume of content decaying even on climate-controlled archive shelves. Unfortunately, too many archives are struggling with high costs, stretched budgets, and a paucity of staff to do anything else. The difficulties of resource allocation are felt acutely in the archival setting. Tapes are neglected because of staff constraints. The New York Public Library, Dance Division, for example, lacks basic condition information for approximately 6,000 of their videotapes. In many archives, the reformatting of their tapes is done piecemeal—and the backlog of tapes is never finished. Similarly, the Theatre Collection at Harvard University has some 5,000 tapes that have not even been inventoried. The problem of tape volume outstripping an archive’s staff resources is evident throughout the field of audiovisual preservation, and it shows no signs of abating.

The traditional method for preserving the content of magnetic media collections is migration (i.e., re-mastering) to new tape stock. Practiced universally by the archive community, migration has been seen as the only solution for aging collections, until recently. Migration has been used for several reasons: format obsolescence, tape degradation, and to create access copies. Formats become obsolete because manufacturers cease to make machines and sell repair parts, and specialists who can maintain such tape players and recorders may no longer be available at that facility. Sony's Umatic format, for example, is going extinct. Sony has stopped manufacturing these playback decks, those that exist are aging, and the knowledge required to maintain them has become scarce and expensive. Formats such as Hi-8, widely used by small dance festivals and companies, are also rapidly being discontinued.

Migration is also necessary when tapes have undergone typical material degradation, from aging, or have been damaged in an accident or disaster, such as a fire. This type of restoration is often the most expensive as it must be done manually by specialists working off-site.

Migration from masters to access copies is common, and it enables archives to share their collections without compromising the safety of their original tapes. In most cases, providing access to rare and valuable content is part of an archive's mission; in the dance community, this approach is critical to the advancement of the field and the education of dancers. Unfortunately, making access copies requires playing back the master, often repeatedly, potentially putting that tape at risk in the long term. Also, if consumer-level equipment is used, access copies can exhibit signs of generation loss; that happens when copying VHS to VHS, with no intervening corrective equipment, such as a time-base corrector.

Archivists, as well as those who fund archives, already understand tape-to-tape migration to be a widely accepted preservation strategy. Typically, when grant-making organizations provide funds for a migration project, the scope of the project is described in numbers of completed tapes. Since

migration is so well understood—either to and from identical formats, or from one format to a different format—there is a reluctance to seek alternatives. Libraries and archives have developed tape-oriented infrastructures; their workflow is geared toward handling cassettes and magnetic tape. Given the history and momentum of tape-to-tape migration, it is not surprising that archives and funders cannot or will not plan for the future preservation of their collections. However, doing nothing, and holding our collective breath, is not an option. The backlog of tapes will continue to degrade, in perpetuity, unless there is significant change.

For the archival field, mass digitization of video as a preservation strategy is a very exciting development. Historically, digitization projects in larger archives have been focused on the creation of low-quality digital files for internal access copies or for use in Web streaming. High-quality, uncompressed or lossless digitization of any footage requires large amounts of hard-drive storage, as well as the accompanying computer equipment and training to use it. Few archives, dance or otherwise, have had the resources to use digitization as a true preservation strategy. Consequently, “lossy” digital formats—those that lose, edit out, or throw away information in the digitizing process—have been the rule.

The seemingly permanent nature of digital distribution media, such as DVDs, has spawned much interest in getting footage off tape and onto something different. If, for example, a dancer’s agent or a dance company requests copies of his or her performance work on DVD, there seems to be little need for the dancer to keep his or her tapes around after spending time and money to have them digitized. The conventional wisdom is that a DVD must be better than tape: they are solid, waterproof, and, according to various marketing campaigns, supposedly able to stand up to worse conditions than tape. On a standard television screen, the picture from a DVD looks good. DVDs are small, lightweight, and easy to carry and to send to anyone who asks, easy to play back at a home or office, and DVDs take up little space on shelves compared to tapes. While manufacturers may claim DVDs and CDs have shelf lives upwards of 100 years, there is much uncer-

tainty about these claims. Recent reports of “DVD and CD rot” are beginning to send ripples of anxiety through the archival community consumers.⁴

Whether or not DVDs are physically archival over the long haul is only one issue. The actual video signal contained therein should be examined for archival quality. Currently, MPEG-2 is the broadcast standard for the digital distribution of video content, used for cable and satellite television transmission, as well as for DVDs. While this form of encoding looks more or less attractive on a standard television screen, whole frames of video are lost, thrown away in the digitizing process to get the file small enough to fit onto the DVD media. Because of this limitation, MPEG-2 does not conform to the Committee’s requirements for a preservation-quality format.

While these encoding standards are in common usage in the broadcast industry, archives have different needs. The loss of any information when remastering is simply not acceptable. When looking ahead to the digitization of a rare collection of videotapes, newer encoding standards must be evaluated.

⁴ <http://www.cnn.com/2004/TECH/ptech/05/06/disc.rot.ap/>

INNOVATIVE IDEAS FOR THE PRESERVATION OF VIDEO

Making the leap from dedicated videotape formats to generic digital files is no small task. There are many factors to consider before dedicating resources and budgets to the digitization of a tape collection, as well as the need for a general re-evaluation of archival workflow.

First, and most obviously, digital files are not tape. While hard drives could be construed as “physical media,” there is a conceptual difference between digital files and magnetic tape. Tape is a linear medium, on which information can be organized in a single, immutable way. Defects in a tape result in errors during playback and migration. Hard drives, on which digital files are stored, should be thought of as nonlinear and mutable; they can be rearranged, altered, moved, and reconfigured electronically, without damaging the underlying content. This is not to say that hard drives are indestructible, far from it—but, they are more systemically flexible than tape.

Transferring a tape collection to digital files requires a completely different set of hardware from a tape-based infrastructure. Tapes are played back on format-specific video decks, such as a Sony Betacam SP deck. Hard drives “live” inside computers; mass digital storage occurs inside arrays of hard drives. While a tape can be played back with simply a video deck and a television, playing and storing video in digital files requires computers. Once you move from one or two video files stored on hard drives into the realm of mass storage (hundreds or thousands of large video files), more complex hardware is required to organize and preserve the content.

In addition to hardware, there are software concerns: operating systems, file organization, security, and backups, to name a few. Advanced hardware tends to require the most recent software available, and specialized hardware must be supported by specialized software. Digitally stored video files are still large and cumbersome, and the computers that move them around

need to be speedy and reliable. Instead of simply having a single video playback deck and TV, to use digital video files, you need a complex system of computer hardware and software working in harmony to achieve the desired results. Also, staff re-training is an inevitable requirement. By giving archive staffs the knowledge they need to use new technology, you enable them and your organization to reap the full benefits of using digital video files. It is important to consider computer knowledge and skill sets when looking forward to future staff hiring in relation to a video digitization project.

All of this new hardware, software, and training results in an archive that appears very different than traditional, tape-based archives. New models in preservation are developing, and the stakeholders in archives and their missions may not immediately grasp the concepts of digital preservation and migration. Executive boards, donors, and grant makers should be included in an archive's transformation to digital file storage, therefore grant-writing efforts need to be updated to reflect the changing systems of video preservation. The long-term advantages and cost savings of digital files are an attractive addition to funding requests.

The evolution from tape-based to digital files has not been rapid. The archival community has been embracing digital technology slowly, and there is much discussion regarding the best way to gain the benefits of "going digital." Uncertainty and confusion regarding the technology cause archives to be hesitant about committing their resources to large-scale digitization projects. This is to be expected; the uncertainty will decline as more successful projects become available online. Digital video technology will become familiar and desirable as it can be accessed from ever increasing archival environments.

When a master recording is digitized, it should be done only once. One of the biggest costs in any digitization project is the transfer from tape to file. Consequently, the file that results from the transformation must be of the highest quality available, since it is inefficient to consider redigitizing in

the future. Further, the cost of re-digitizing becomes less attractive when you consider that, during the time that has passed from your first digitization, the original tape master will have further degraded, and the playback decks and associated gear will have further aged. The goal of any digitization project should be to create the best possible file from the footage, in light of the fact that subsequent digitization may not be possible because of cost and media aging.

THE DETERMINATION AND SPECIFICATIONS OF PRESERVATION FILE FORMAT CANDIDATES

Based on discovery findings in Phase I, we began to determine and specify file formats that appeared to be good candidates as a “Preservation File Format.” Based on early interim results from Phase I, both AAF and MXF file wrapper types appeared to be good candidates for consideration. What eventually became apparent, however, was the necessity to break this module into two sections. The first focus was to examine video compression technology in depth, specifically to scrutinize both lossy and lossless compressed file types. Second, we examined so-called *file wrappers*. The most important consideration quickly became the determination and specification of the video compression technology and technique, most any of which could be contained in the chosen wrapper format. Thus, the video compression technique determination, specification, and testing quickly became the main challenge of the Phase II project.

Two classes of video compression technology were examined for Phase II: lossless compression and lossy compression. These basically different systems of compression both have the capability of producing very good quality imagery, but each differs dramatically in terms of how it gets there and the compromises taken in the compression process.

LOSSLESS COMPRESSION

Mathematically lossless compression technology (referred to here as *lossless*) is the technology familiar to those who are accustomed to using computers. In fact, for many years, lossless compression technology was the only compression technology that existed and was used in the data processing and computer fields. Essentially, lossless compression techniques make a file smaller for storage purposes, without any change to the content of the file. (That is to say, the file before compression is identical to the file after reinflation.) There are many different techniques for lossless compression, and one is the familiar .zip file, frequently used in the Microsoft Windows Operating System environment. When taking a file and “zipping it,” the resultant file is usually smaller than the original file. When accessing it again at a later time, it is brought back to its original identical size. Zip files are just one example of literally hundreds of techniques that can be used to compress data for storage purposes, while keeping the content intact.

Lossless compression techniques work in many different ways, and many of them use complex mathematical techniques to optimize the results, but an easy way to understand them is to consider a technique called “run length encoding.” In run length encoding, we compress simple redundancies merely by changing notation. For example, one can store a series of twenty-one of the letter A as follows:

AAAAAAAAAAAAAAAAAAAAA.

Another way to do this would be to simply store as:

21A.

In this case we have reduced the storage space from twenty-one characters to three. The ratio of compression in this example is 21:3 or, when reduced, is generally discussed in the industry as a 7:1 compression ratio. The higher the ratio, the smaller the amount of storage space is required for any given

amount of information and the more efficient a technique is. Compression ratios quickly can be thought of in terms of cost. If one has to pay \$1 per gigabyte (GB) of storage, then it would cost \$100 to store 100 gigabytes of uncompressed data. If this data can be compressed at a ratio of 100:1, then this same data can be stored at a cost of only \$1.

If such large cost savings are possible, and the results are identical, why not compress everything? The maxim “There is no free lunch” applies well to compression in general and to lossless compression in particular. It takes time and processing power to compress data; in particular, lossless compression requires computer power to perform, and this often equates to time. In some cases there is no problem with waiting a period of time to compress information, but in other applications it is highly undesirable and impractical. In many real time applications, such as video, information must be available at certain, very tight time intervals in order to make a properly synchronized picture. If the information takes too long to compress or decompress, the results can be disastrous, often resulting in a damaged picture or file—or no picture or file at all.

Lossless compression in the video application area has had two major problems. First, too much processing power has been required to allow for compression of so much data in real time. Second, the fact is that the compression ratios are fairly inefficient (the ratio is much poorer than for lossy compression techniques). This inefficiency directly relates to storage cost, which is always a very important issue for preservation purposes, and is especially so in the dance community, where funding for it has usually been limited. What must be considered, however, is the tremendous advantage of having identical information before and after compression—perhaps the key requirement for preservation purposes.

During the course of this study, we discovered that a new standard was being developed for video compression. This was no surprise, because there are many different standards in existence and several standards in development. What was of particular interest with the new JPEG2000 standard is

that, in one section of the standard, there is an option for lossless compression. This truly is a first. While there were discussions of lossless compression in other standards, as a practical matter this was “for real.” In addition, the technique promised to be mathematically lossless (other techniques have been called lossless but, in fact, were only “visually lossless,” which is to say mathematically lossy, and therefore in the other class of compression techniques). This new JPEG2000 standard promised to be of enough interest to the broadcast community that dedicated hardware would be produced, thus allowing both the compression and decompression to occur in real time.

The JPEG2000 standard allows for an uncompressed file to be displayed with various levels of quality. This is a very important element for the dance community, because it means that one does not have to keep several versions of files or different versions for different applications (archival storage versus remote viewing, for example). It is possible to produce copies at lower resolution and bit rate for some applications, while keeping the original file intact and losslessly compressed.

Finally, the storage ratio of approximately 3:1 was not spectacular, but it was significant enough when compared to uncompressed files to warrant serious consideration. JPEG2000 quickly was added to our selection list for experimentation purposes.

LOSSY COMPRESSION

Unlike mathematically lossless compression, lossy compression is a technique whereby the original file that is compressed differs from the resultant inflated file. The reason it is called *lossy* is because some of the information is in fact lost. These techniques are fairly new because they are of limited utility to most data-processing applications. By definition, most applications require both the original and copy to be identical, and for these applications, lossy compression is unsuitable. In transferring video, however, lossy compression techniques hope to fool our eyes, by presenting pictures of “good enough” quality that we may not be able to see the difference.

Lossy compression techniques do have several advantages. First, lossy compression is a fact of life in the video industry, where a great deal of the modern equipment records by using lossy compression technology. Most “born digital” recordings created today by consumer or prosumer equipment are already compressed. Lossy compression is a fact of life in the video world, and one cannot avoid it. Because of its acceptance in the marketplace, there are a wide variety of techniques from which to choose. In addition to having a wide variety of techniques and standards, there are also issues of *bit rates*. From a practical point of view, this means that one can use the same technique and “tune” it in terms of quality. Higher quality inevitably means a higher bit rate and a lower compression ratio and, therefore, a higher cost. Lower bit rates can be distributed through channels with limited bandwidth. For example, video streaming can occur over relatively slow or limited bandwidth systems, such as telephone modems, to produce moving pictures. This is not possible with higher bit rate systems or systems that inherently need more bandwidth, such as lossless compression. Therefore, it became necessary to consider the distinction between signal distribution and archival storage.

For the purposes of our experiment we chose to test several lossy compression techniques at several different bit rates. It was important to find out how good or how poor the images really were, and—in fact, if it might be possible to have a preservation strategy that is “good enough” to accomplish the several different requirements of the dance community.

Finally, lossy compression can be extremely efficient. Unlike lossless compression, where ratios of only about 3:1 are possible, lossy compression has reasonably high quality results with ratios of 40:1. In order to test lossy compression techniques and their suitability for dance footage, it was important to test the algorithms at different bit rates on diverse types of materials to see if they responded differently to material that was visually dissimilar. In short, we tried to answer the question “Do different lossy compression techniques at different bit rates produce different results with different visual material?”

FILE WRAPPERS

When making the move from audiovisual records contained on videotape to audiovisual records contained in digital files, we face a number of choices when deciding on a destination format. The essential elements are a high quality (preferably lossless) video and audio recording process (or algorithm), and a means by which detailed data about the media content can be linked and preserved in the digital file. Audiovisual digital media frequently makes use of the concept of “file wrappers,” which generally combine video files, audio files, and metadata into a single, unified format.

File wrappers can serve as “codec wrappers,” a generic video file format that simplifies the play back of various codecs (“code/decode” packages). It allows an operating system to select the proper codec locally or to find it on a network or Internet resource. Examples of “codec wrappers” are the AVI format and Apple’s QuickTime, which act as the interface and container for the digital media file(s).

With the increased importance of metadata files in both the preservation and production industries, a number of rich metadata-supporting file wrappers have emerged over the past few years. The leading candidates for advanced metadata handling are Media Exchange Format (MXF) and the Advanced Authoring Format (AAF). These two formats allow program content or essence, such as video and audio, to be wrapped in a file in a structured and standard way, along with its metadata. However, these standards differ in their intended applications.

AAF

The Advanced Authoring Format (AAF) is a professional file interchange format designed for the post-production and authoring environment. AAF solves the problem of multivendor, cross-platform interoperability for computer-based digital production. AAF does a number of things:

- (1) it allows complex relationships to be described in terms of an object model;
- (2) it facilitates the interchange of metadata and/or program content;
- (3) it provides a way to track the history of a piece of program content from its source elements through final production;
- (4) it makes it possible to render downstream (with appropriate equipment); and
- (5) it provides a convenient way to “wrap” all elements of a project together for archiving.

By preserving comprehensive source referencing, and abstracting the creative decisions that are made, AAF improves workflow and simplifies project management. (AAF Association, <http://www.aafassociation.org/>)

AAF was introduced in 1998, promoted by the leading companies in their respective fields: Avid for video editing and Microsoft for digital media. AAF originated with Avid’s Open Media Framework Interface (OMFI), which was then further developed by Microsoft. The AAF Association now consists of many prominent companies in the converged video/digital media field, such as Adobe, BBC, Discreet, Pinnacle, and several others.

AAF is intended as a vendor-neutral architecture to support a variety of nonvideo advanced media types, such as text files (including HTML and XML objects), plus 2D and 3D objects. It serves as a container for media and its associated metadata, with emphasis on compositional metadata, describing how clips are comprised, edited, arranged, and modified, as well as a record of “versioning,” a history of changes made to the associated media file. Examples of integrated compositional metadata include edit decision lists (EDLs), which originated with linear editing but can be integrated effectively into AAF files. More advanced structures, such as AES-31 and OMFI, are also superseded by AAF, with some vendors offering translation/upgrade tools.

The AAF format separates the editorial information from the media source, to enable the exchange of essence as well as its associated metadata in one standard. This file contains a collection of data that includes an index of all objects within it: the metadata objects, the dictionary that defines those objects, and optionally, the essence media itself. Within the “material object” categories of metadata are the following:

- identification and location (how the item is uniquely identified)
- administration (rights, access, encryption, security, etc.)
- interpretive (names, artists, etc.)
- parametric (signal coding and device characteristics)
- process (editing and compositing data)
- relational (describes the relation between various pieces of metadata and/or essence—in effect the “verbs” in the equation)
- spatio-temporal (places, times, things, camera angles, etc.).

AAF was designed for production environments, largely as an authoring tool, and is not intended as an end-user delivery or presentation format. The format was designed to provide a standard for production and post-production workflows, where the convergence of multiple nonvideo media types would most benefit from a standard packaging. These production environments are typically required to combine multiple-input source types from several production facilities. At the same time, the production industry is in the latter stages of transitioning from analog sources and physical media to network-based digital media, which this standard addresses. AAF was designed to standardize the development process and provide more efficiency in these collaborations.

AAF has also been designed to be a flexible format, with support for “private” metadata, which would allow certain vendors to collaborate with a particular set of metadata for their own particular process. While defined and promoted by Avid and Microsoft, AAF is an open standard, not owned or controlled by a single company. It is developed on the Source Forge open source platform (www.sourceforge.net/projects/aaf).

Using AAF, the metadata may also be separated from the original essence audiovisual content; in addition, the file wrapper may make use of external references to the original material.

MXF

MXF is the acronym for Material Exchange Format, an open file standard designed for the interchange of audio-visual material with associated data and metadata. MXF is a file format for the exchange of program material between and among servers, tape streamers, and digital archives. Its contents may be a complete program as well as complete packages or sequences. There are basic facilities available for cuts between sequences and audio cross-fades; this way the sequences can be assembled into programs. MXF is self-contained, holding complete content without any need of external material.

MXF bundles together video, audio, and program data, such as text—together termed *essence*—along with metadata, then places them into a wrapper. Its body is stream-based and carries the essence and some of the metadata. It holds a sequence of video frames, each complete with associated audio, and data essence, plus frame-based metadata. The latter typically comprises time code and file format information for each of the video frames. This arrangement is also known as an *interleaved media file*. MXF was implemented to improve file-based interoperability between servers, workstations, and other content creation devices in a networked facility. (The PRO-MPEG group, <http://www.pro-mpeg.org/index3.html>)

MXF defines the data structure for the audio and visual material (essence) plus associated metadata. This metadata is defined in a header and footer. The header and footer generally contain sections for “partition” (the structure of sections and essence containers), “metadata” (containing structural and descriptive information about the essence), and the “index” (which provides for instantly accessing points of essence in the file).

Technically, the MXF format is a subset of AAF, designed for more efficient, linear essence. As with AAF, MXF is an open standard. MXF's metadata structure is designed to cover descriptive metadata and structural metadata, including both information about the media essence and synchronized events. MXF lends itself to random-access searching based on this metadata.

MXF provides for well-defined "packages" within the metadata that allows for easy translation from certain editing structures, such as an EDL and external references to original source material. For example, the Material package is the final timeline for output and use by an end-user on a one-time track; the File package lists all clips, with respective time-codes listed in order; the Source package contains pointers to actual essence files. Within the given structure of these packages, the MXF user has quite a bit of flexibility in defining a metadata schema for a particular file or series of files.

MXF is not specific to any compression scheme. It supports MPEG and DV, uncompressed, and is open to support future compression technologies. It has widespread industry support and has been offered as a published, open standard.

MXF VS. AAF

When considering a digital-media wrapper format for archival purposes, MXF and AAF offer many features to augment and extend the value of the contained video and audio record. Both file wrappers have the flexibility to wrap high-end uncompressed digital media, as well as lossless compressed media, such as Motion JPEG2000.

Both MXF and AAF are container formats, and can be considered complementary technologies to the production industry. MXF is not designed to be a composition format; instead it provides a useful container to associate

media and a standard set of metadata. AAF carries compositional information useful for the production and post-production process as related to the creation or modification of the media file, while MXF is better suited to carry information about the media itself.

One issue with AAF is that the edit lists and other process metadata may be of a proprietary or sensitive nature, since it may represent unique or otherwise privileged information on how a piece was created or modified. While that information enhances and expedites workflow in production, it serves no value to the end-user.

Another distinction between AAF and MXF is in the location of source material: while AAF may contain pointers to essence contained outside the file, MXF must contain essence files within the MXF file—and must not require access to outside material.

Therefore, MXF is well suited as a candidate for both preservation and access of archival audiovisual content and records, based on the broad adoption of the standard, the flexibility to contain detailed content metadata, a structure designed for end-users, the requirement to have media files included in the wrapper, and its support for lossless compressed media.

CONSTRUCT THE SOFTWARE (IF NECESSARY) TO CREATE PRESERVATION FILE FORMAT CANDIDATES

For the tests, Media Matters assembled several different compression techniques, at different compression ratios, to make preservation file candidates. Since both AAF and MXF file wrappers are capable of containing a wide variety of file types, and both are industry standards and reasonably open, they both pass the test of basic suitability as a preservation file container. What was unknown was the level of industry adoption of each system. When starting the study, we gave a very optimistic assessment of adoption—and, frankly, expected to see industry-wide adoption of both wrapper systems by the end of the study. Unfortunately, this is not the case; behind the press releases is the sad fact that real-world adoption has been slower than anticipated. It does appear that MXF has some industry support, with several manufacturers promoting it. As an example, at the 2004 National Association of Broadcasters show in Las Vegas, Snell and Wilcox, a fairly large company that produces video post-production equipment, announced that they were “giving away” software that allowed the making of MXF file wrappers.⁵

⁵ <http://www.postmagazine.com/post/article/articleDetail.jsp?id=87277>

PRODUCE A FOOTAGE TEST TO INCLUDE DANCE FOOTAGE AND OTHER TEST FOOTAGE

While we are disappointed at the speed of wide industry deployment of these wrapper systems, the reality is that both AAF and MXF are reversible by design. This means that archives could, in fact, choose to adopt either format and be secure in knowing that they can extract the essence and metadata if these standards are not widely accepted and another wrapper system develops. Because of this reality, from a testing point of view, we decided to concentrate on the compression technology—which we believe is the major exploration issue—no matter which wrapper format is chosen. The choice of either wrapper should have virtually no effect on the visual quality of the stored imagery. By contrast, the compression technology has a huge effect on the visual quality and, therefore, also on the preservation of the content.

While compiling research for the design of the test, we believed that it would be necessary to use test footage, other than dance footage, in order to determine values for subjective quality analysis. The Sarnoff Laboratory's JNDmetrix IQ tools require the use of specific test footage, which has nothing to do with dance footage, but is electronic test footage designed to test encoding systems. This type of system is called a Full Reference system (FR), and while useful for some applications, it was less than optimal for us. At the time of the initial proposal, it was the only option in the marketplace. Fortunately, we were able to find a vendor that uses absolute, or Non-Reference (NR), analysis. Using this newer approach, we were able to concentrate on the specific analysis of dance footage rather than test patterns. While test patterns are useful for technical analysis, we were much more concerned about the actual performance of compression algorithms on real-world footage, which has been limited in the past because of the lack of NR tools. Our test footage therefore was solely dance footage, and the new NR software allowed us to obtain more useful information than anticipated.

METHODOLOGY

Samples of dance video files were chosen with assistance from the New York Public Library (NYPL) and Jacob's Pillow, representing a variety of styles of dance shot on a variety of videotape formats. The chart below (Figure 1) outlines the clips that were used, where they came from, and on which format originally recorded.

FIGURE 1

<i>Source</i>	<i>Choreographer/ Performers</i>	<i>Work/Location/Date</i>	<i>Format</i>
NYPL Clip 1	Concept and Choreography by Elizabeth Streb Performed by Streb/Ringside	<i>Bounce</i> Excerpt from Streb Joyce Theater, New York City 19 December 1997	Betacam SP
NYPL Clip 2	Concept and Choreography by Elizabeth Streb Performed by Hope Clark	<i>Breakthru</i> Excerpt from Streb Joyce Theater, New York City 19 December 1997	Betacam SP
NYPL Clip 3	Concept, Direction, and Choreography by John Kelly Performed by John Kelly	<i>Pass the Blutwurst, Bitte</i> Excerpt La MaMa E.T.C., New York City 12 January 1995	3/4" Umatic
NYPL Clip 4	Mar Gueye and N'Geuwel Sabar Dance from Senegal Mar Gueye, Company Leader and Choreographer	<i>Domba Concert of Dance</i> Excerpt from Niani Badenya, The Mandeng Heritage Heckscher Theater of El Museo del Barrio, New York City 1 June 1997	Betacam SP
NYPL Clip 5	Conceived, Choreographed, and Directed by Ralph Lemon	Excerpt from <i>Geography</i> Yale Repertory Theatre, New Haven, Connecticut 4 November 1997	Betacam SP
NYPL Clip 6	Danced by Cok Ratih Iriani and Made Lila Arsana	<i>Oleg Tambulilingan</i> or <i>Bumblebee Dance</i> Excerpt from The Dancers and Musicians of Bali Town Hall, New York City 22 March 1996	Betacam SP

Methodology

FIGURE 1 (CONTINUED)

<i>Source</i>	<i>Choreographer/ Performers</i>	<i>Work/Location/Date</i>	<i>Format</i>
NYPL Clip 7	Danced by Savion Glover and Gregory Hines	<i>Improvisation</i> Excerpt from Tap City New York City Tap Festival 2001 New 42nd Street Theater 12 July 2001	Betacam SP
NYPL Clip 8	Created and Performed by Basil Twist	<i>Primo Ballerino Stickman</i> Excerpt from Deaths and Entrances Mother, New York City 4 November 1998 Martha@Mother with Richard Move	Betacam SP
NYPL Clip 9	Choreography by Dwight Rhoden Artistic Direction by Dwight Rhoden and Desmond Richardson	<i>Inkblot</i> Excerpt from Complexions— A Concept in Dance Brooklyn Academy of Music Majestic Theater 19 September 1997	Betacam SP
NYPL Clip 10	Directed by Francisco Nevarez Burgueno	<i>Estampas y Tradiciones</i> Excerpt from Mano a Mano, Cultura Mexicana sin Fronteras Haft Auditorium, Fashion Institute of Technology, New York City 16 December 2001	Betacam SP
NYPL Clip 11	Artistic Direction by Erwin Kilip Performed by Bibak	<i>Bendiyan</i> Thanksgiving dance, originally of the Ibalois Tribe of Benguet Excerpt from Pagbubunyi: A Celebration of Filipino Culture and Heritage Washington Irving High School, New York City 2 April 2002	Betacam SP
NYPL Clip 12	Choreography by Tyler Walters Carolina Ballet Artistic Director, Robert Weiss	<i>While Going Forward</i> Excerpt A.J. Fletcher Opera Theater Raleigh, North Carolina 19 May 2001	Betacam SP
NYPL Clip 13	Created by Amy Sue Rosen and Derek Bernstein Danced by Sally Bomer, Victoria Boomsma, Thom Fogarty, Sam Keany, and Phillip Karg	<i>Abandoning Hope</i> Excerpt from Triage The Duke on 42nd St., New York City 17 March 2001	Betacam SP

Methodology

FIGURE 1 (CONTINUED)

<i>Source</i>	<i>Choreographer/ Performers</i>	<i>Work/Location/Date</i>	<i>Format</i>
NYPL Clip 14	Choreography by David Parsons Dallas Black Dance Theatre Founder and Artistic Director, Ann Williams	<i>Nascimento</i> Excerpt from <i>Dance Women/ Living Legends</i> Aaron Davis Hall, City College, New York City 15 November 1997	Betacam SP
NYPL Clip 15	Cathy Weis Projects, Nova Productions from Skopje, Macedonia	<i>Not so Fast, Kid!</i> Excerpt from <i>Show Me The Kitchen</i> , New York City 11 January 2001	DVCAM
NYPL Clip 16	Choreography and Text by Neil Greenberg Performed by Ellen Barnaby, Christopher Batenhorst, Neil Greenberg, Justine Lynch, and Jo McKendry	<i>Not-About-AIDS-Dance</i> Excerpt The Kitchen, New York City 15 December 1994	3/4" Umatic
NYPL Clip 17	Period Choreography by Catherine Turocy New York Baroque Dance Company Artistic Director, Catherine Turocy	<i>Menuet à Quatre</i> Excerpt from <i>Soirée Baroque en Haïti</i> Florence Gould Hall, New York City 2 November 2003	DVCAM
NYPL Clip 18	Choreography by Marcea Daiter Dallas Black Dance Theatre Founder and Artistic Director, Ann Williams	<i>Vodun Zépaule</i> Excerpt from <i>Soirée Baroque en Haïti</i> Florence Gould Hall, New York City 2 November 2003	DVCAM
Jacob's Pillow Clip 19		<i>Chore</i> Student Showing 25 June 1992	Hi-8
Jacob's Pillow Clip 20		1992 Gala Ted Shawn Theatre	Hi-8
Jacob's Pillow Clip 21	Choreography by Trisha Brown	<i>Informance</i> Ted Shawn Theatre 10 August 1986	VHS
Jacob's Pillow Clip 22		<i>Halau Hula O Hoakalei</i> <i>Ka Pa Hula Hawai'i Hula</i> Excerpt from performance 3 August 1989 and workshop 4 August 1989	VHS

Each clip was selected for the type of video content contained, with special attention also paid to imagery known to be problematic when digitally compressed.

The original VHS, Hi-8, Umatic, Betacam, and Betacam SP dance footage was copied to two Betacam SP tapes. This tape was then encoded as raw, uncompressed digital data as .avi files. The AVI files were created, using playback from Sony UVW-1800 Betacam SP, Sony DSR-30 DVCAM, Sony EVC100 Hi-8, and JVC BRS822U SVHS Decks. The signal was analyzed and levels were set using an OmniTech OmniView Video Analyzer. The analog signal was fed into a Digital Rapids StreamZ 1500 for uncompressed capture via a Leitch DPS-290 Time Base Corrector/Synchronizer.

Dance footage originating from DVCAM was captured as raw DV signal data directly to a computer from a DVCAM deck. The raw digital data and digital formats were processed by software and compressed with commonly used compression algorithms at a variety of generally used bit rates. The result of this approach was a single uncompressed file type that could be compressed using the various algorithms in a controlled fashion. These files could be compared to the original uncompressed AVI files.

COMPRESSION

The experiment compared the results of reformatting the test footage as uncompressed video, lossless compression, and higher-end and lower-end lossy compression. (Examples of lossy compression include DV25, DV50, and the MPEG2 long group of pictures [long GOP] at data rates of 50 to 100 megabits per second.)

The uncompressed AVI files were processed using Discreet Cleaner XL and Discreet Cleaner 6, using the following codecs:

.mov files =

- Sorenson video 3,
- 640 x 480 millions of colors
- 29.97 fps
- Interlaced bottom field first
- Key frame every 300 frames
- aspect ratio 4:3
- bit rate limit 1200 kbps
- spatial quality 50
- image smoothing on

.mp4 files =

- MPEG-4 Video,
- 640 x 480 millions of colors
- 29.97 fps
- Interlaced bottom field first
- Key frame every 300 frames
- aspect ratio 4:3
- bit rate 1229 kbps

.rm files =

RealMedia 9
640 x 480 millions of colors
bit rate 1067 kbps constant bit rate
29.97 fps
4:3 aspect ratio
progressive (no option for interlaced)
Key frame every 300 frames

.wmv files =

Windows Media Video 9 Professional
bit rate ~1340 variable bit rate
29.97 fps
4:3 aspect ratio
Interlaced bottom field first
Key frame interval 300 frames

mpeg-2 =

20 Megabit 640 x 480
29.97 fps
Interlaced bottom field first
constant bit rate
4:3 aspect
GOP Pattern IPBBIPBB
Long GOP
Sequence headers for each GOP
High Motion Search Range

jpeg2000 =

Motion JPEG2000 Kakadu

variable bit rate (lossless)

29.97 fps

Interlaced bottom field first

4:3 aspect ratio

5/3 Reversible

millions of colors

Bit rates were chosen, based on common types for each codec. A slight variation in the bit rates was due to the varying bit rates of the accompanying audio tracks that are most often found with each respective codec. The exceptions to this are the two high bit-rate codecs, MPEG-2 and JPEG2000.

The raw, uncompressed sample clips were run and analyzed by the software. This established a baseline to compare the compressed sample clips. Next, the compressed clips were run through the same analysis software. The results from the raw analysis and compressed analysis were compared and the output of the analysis metrics was expressed graphically. Conclusions were drawn, based on this output as to where in the signal—and to what extent—compression algorithms created acceptable or unacceptable levels of loss of quality.

Compressed clips were then watched and compared to the raw clips and the visual confirmation of the software was confirmed. Conclusions were made based on both software analytics and human perceptual confirmation.

CODEC ANALYSIS

MJPEG2K

For testing purposes, motion JPEG2000 (MJPEG2k) was selected for its intrinsic and robust support for lossless compression, a feature of particular need to archivists. Motion JPEG2000, is a video adaptation of the new JPEG2k standard for still photos. It treats a video stream as a series of still photos, with each video frame compressed separately using, the JPEG2k still image compression standard. No interframe compression means that no frame differencing or motion estimation is used to compress the images, which makes it ideal for frame accurate editing without any loss of image quality.

MPEG-2

MPEG-2 was selected for our testing because of its widespread use in industrial distribution video systems, as well as its nearly ubiquitous use in consumer DVD formats. The MPEG-1 international standard for video compression of audiovisual signals was originally designed for CD-based applications that maxed out at roughly 1.5 Mbits. Its successor, MPEG-2, supports the higher bit rates utilized by broadcast applications, as well as support for progressive and interlaced display technologies, such as computer monitors and televisions. The full MPEG-2 standard defines various “profiles” for its different implementations that use different algorithms and toolsets. It provides compression schemes both intraframe (within a frame) and interframe (between frames): these are Discrete Cosine Transform (DCT) encoding and motion-compensated frame prediction, respectively. However, these schemes may introduce patterns of loss in the original data.

MPEG-4

Another test choice was an advance in the MPEG family, MPEG-4. It expands video delivery systems into new multimedia applications such as video conferencing and Internet video streaming. It addresses key issues—added robustness across potentially unreliable networks—such as the Internet or wireless mobile networks, so the end-user experience would be as seamless as possible. MPEG-4 allows for a new level of interactive functionality, so that in addition to strictly audio and video content, an author can include titling, animations, and other multimedia content. Since it was designed with computer networks in mind, it also has better support for high-quality decoding through very low bit rates, such as the sub-56k streams available on telephone modem connections. We chose MPEG-4 because the standard supports the combination of video with innovative computer-based graphics applications and network distribution possibilities. The standard is based on QuickTime.

WINDOWS MEDIA

The .wmv files are Windows Media 9 files, a format developed by Microsoft Corporation, primarily with the goal of streaming video to a large number of viewers. The codec is integrated into Windows operating systems and is also available for Macintosh and other operating systems. We chose it because of its widespread availability and because it is one of the major compression codecs used in the consumer marketplace for the distribution of content. While we understood its limitations in the production of extremely high-quality output, it is of particular importance since it is supported by Microsoft, the clear leader in the personal computer arena. Windows Media is a lossy codec that in its latest incarnation uses a Microsoft-developed (and therefore proprietary) implementation of MPEG-4.

REALMEDIA

RealNetworks was an early pioneer of streaming media over the Internet, with the first widespread commercial success in this area. We selected it for testing because of its widespread adoption of the player on consumer computers. The format has shown consistently improving compression schemes with each version, with the focus on improving quality at the encoding side and allowing backward compatibility with previous decoders too (such as allowing Real9 players to play Real10 content).

QUICKTIME/SORENSEN 3

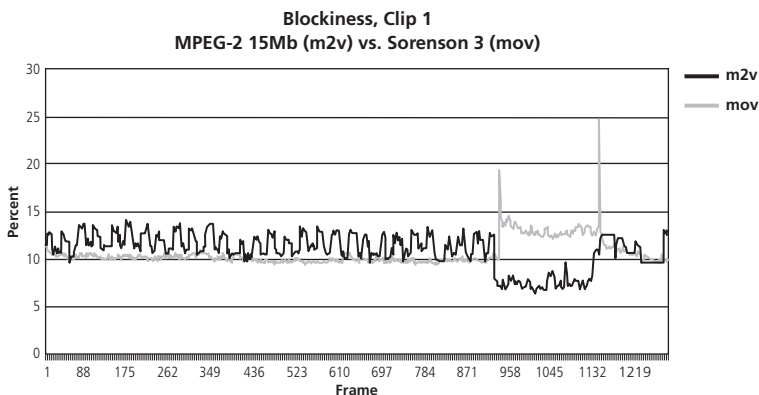
Sorenson 3 is the third-generation codec built by Sorenson Media designed to showcase QuickTime's excellent quality at high bit rates. Among the reasons it was selected for our testing is that it was chosen by Apple Computer for their high-quality online "Trailer Park" section of their QuickTime Website, and it has become a very popular choice for high-end downloaded video on the Web.

THE ANALYSIS OF THE TESTS RUN ON THE FOOTAGE

Media Matters used the Genista software, along with the clips provided by the Dance Heritage Coalition, to perform what might be described best as an exhaustive analysis. Genista software results are unfortunately not graphical, but rather they provide a value for each frame and for each parameter tested. This analysis generated well over 4 million discrete test results on the twenty-two clips that were tested.

While having the values is important, that much data in non-visual form makes it extremely difficult to draw conclusions. We chose to illustrate the Genista core analysis by generating graphs for each parameter measured for each clip. The results were several hundred graphs, which were included in the original version of this report, delivered to The Andrew W. Mellon Foundation in June 2004.

When viewing these graphs, we were especially interested in finding relationships between the job that different codecs performed on the same footage, as well as the reaction of the codecs to differing types of visual images that were occurring in the original. We chose to do a further stage of analysis, presented here, where we illustrate some of the interesting results of the tests. For each clip, we demonstrate some of the interesting relationships graphically and our interpretation of them.



CLIP 1 *Bounce*

Excerpt from STREB, Joyce Theater, 19 December 1997

Concept and Choreography by Elizabeth Streb

Performed by STREB/Ringside

Videotaped by Video D Studios

Excerpt copied from Betacam SP

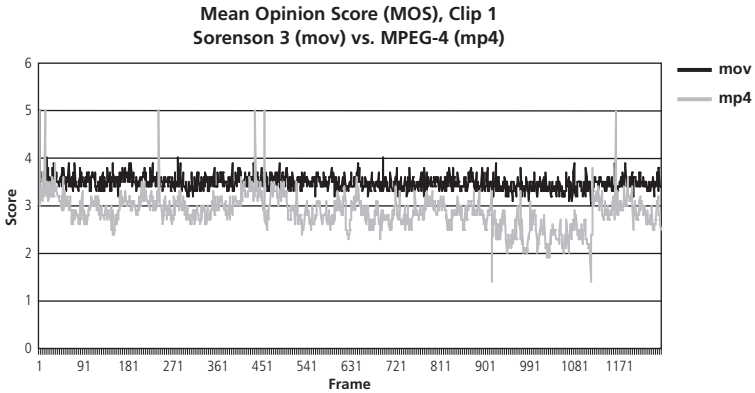
Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

High contrast, multiple dancers, lit for stage and not camera. We liked how performers were entering center area quickly and then exiting. We anticipated a lot of jerkiness and breaks along the lines of their bodies.

In this clip, the Sorenson 3 codec is dealing better with the high motion throughout most of the clip and preventing the clip from becoming exceedingly blocky, especially in the center of the frame, where the dancers enter and exit the overexposed space very quickly. However, when the video cuts to a different camera at frame 917, then back to the first camera at frame 1118, the tables turn and MPEG-2 appears to be dealing better with the cut and overexposure in the center of the frame.

For these experiments we used Genista's Media Optimacy to compare and

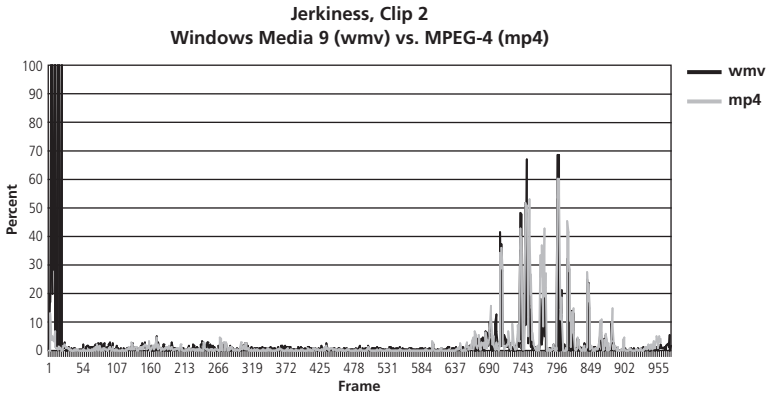


analyze the compressed footage against the original uncompressed footage. One of the key metrics used to summarize overall signal quality is MOS, or Mean Opinion Score. Genista describes this metric as follows:

MOS Prediction: MOS is the Mean Opinion Score obtained from experiments with human subjects. Genista's MOS predictions are metrics that correlate with human perception of video quality and thus with the output of subjective test results....

A set of subjective test data has been used to confirm the high correlation that this measure has with MOS from subjective tests. It should be noted that the accuracy with which this metric reproduces subjective MOS is necessarily dependent upon the type of content used. It has been demonstrated that for typical video content, covering a wide range of motion and texture ranges as well as common PC video codecs, the correlation of the metric with subjective MOS is significantly higher than PSNR.

In this MOS analysis, Sorenson 3 (.mov) delivered consistently better performance in a tighter range than the .mp4 clip. Note, however, that at times the MPEG-4 (.mp4) produced moments of extremely high subjective quality although the average was much lower. By contrast, the Sorenson delivered a more even and better level of quality, although clearly the results are not overwhelmingly good.



CLIP 2 *Breakthru*

Excerpt from STREB, Joyce Theater, 19 December 1997

Concept and Choreography by Elizabeth Streb

Performed by Hope Clark

Videotaped by Dennis Diamond of Video D Studios

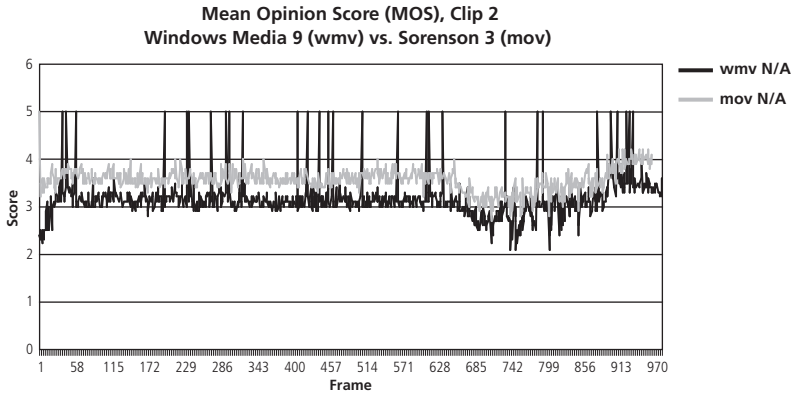
Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

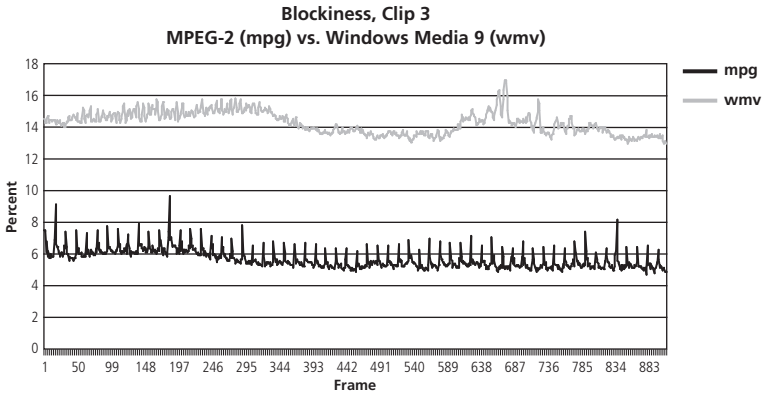
Why we were interested in this clip:

Fast motion of the dancer in the center. The shiny sugar-glass window crashing on impact with the dancer could produce some interesting effects—if compression was high enough, the viewer might even completely miss that.

In the first 20 frames, the camera zooms in abruptly. Windows Media 9 becomes much jerkier, while Sorenson 3 handles this transition more easily. Both codecs have similar difficulty dealing with the motion of the performer as she jumps through the sugar-glass window. This is evident by the relative stillness seen in the video, which correlates to the relative smoothness of the graph from frame 21 to approximately frame 660.



Windows Media performed quite well, considering its lower bit rates, and the efficiency is quite clear for imagery with little movement. It is unclear why Sorenson had such positive quality spikes, other than the possibility that the high-quality spikes are, in fact, not interpolated frames but B frames, which would explain the higher level of quality.



CLIP 3 *Pass The Blutwurst, Bitte*

Excerpt

John Kelly and Company

La MaMa E.T.C., 12 January 1995

Concept, Direction, and Choreography by John Kelly

Performed by John Kelly

Videotaped by Penny Ward Video

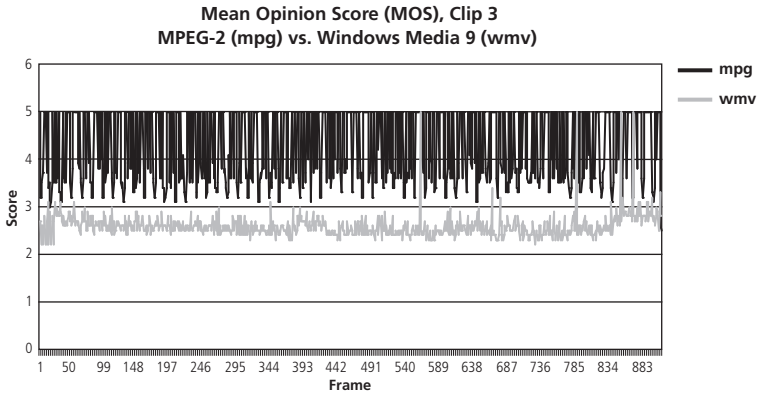
Excerpt copied from 3/4" Umatic

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

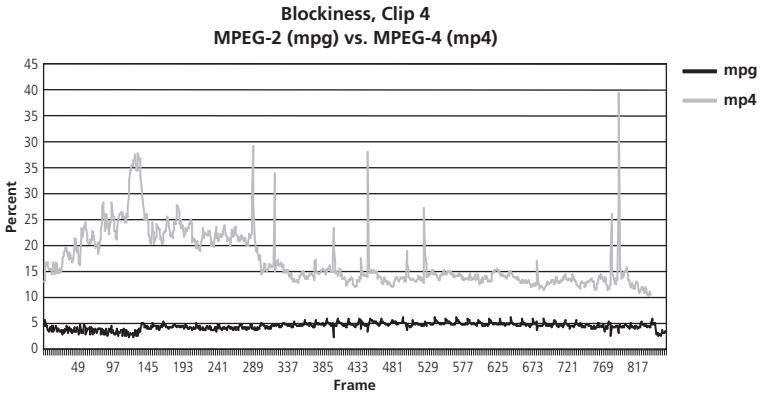
High contrast between dancer and the white board he is holding. Hard edges could become jerky or blurred. Notice the tall shadow the dancer casts in the background—in the .avi it is easier to see, but we thought that as compression increased the shadow would simply disappear into the dark background.

The jagged saw tooth pattern in MPEG-2 data correlates to the performer spinning around, while he holds the white card above his head. The Windows Media 9 data indicates a decrease in blockiness after the performer drops the card and it fades into the low light of the background. Blockiness



increases in Windows Media 9 with increased camera movement, as well as when the camera zooms in and out toward the end of the clip. MPEG-2 seems to be handling those camera changes very well.

MPEG-2 consistently does a better job than Windows Media, but there is tremendous variation in quality during the piece. While Windows Media has consistently poorer results, the consistency may in fact be less distracting to the viewer.



CLIP 4 Domba Concert of Dance

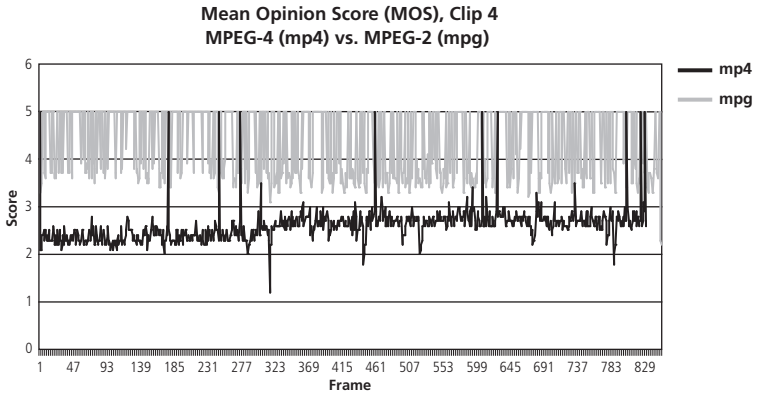
Excerpt from Niani Badenya, The Mandeng Heritage
Heckscher Theater of El Museo del Barrio, 1 June 1997
Mar Gueye and N'Geuwel Sabar Dance from Senegal
Company Leader and Choreographer, Mar Gueye
Videotaped by Mamadou Niang of NextMedia.tv
Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

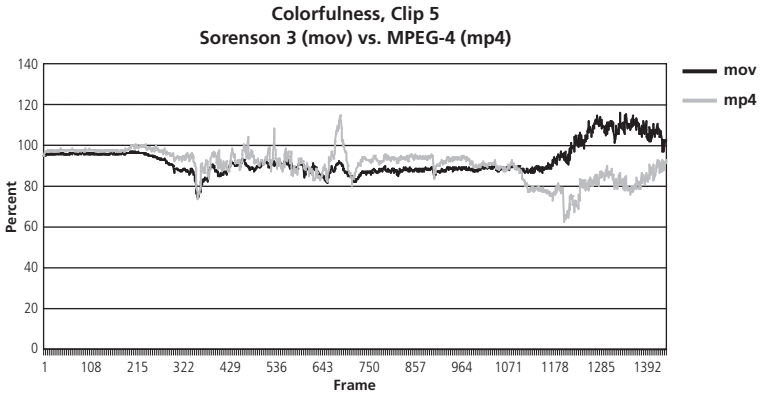
Why we were interested in this clip:

Colorful costumes: have a full color range, and those patterns could easily get lost. Also, the superfast dance steps could get blurry and jerky.

This clip contains very fast movements, multiple dancers, multiple cameras, as well as colorful swirling costumes. The data suggest that 20Mbit MPEG-2 will do much better than the lower bit rate of MPEG-4. Obviously, in most cases, a higher bit rate will produce a better result. The comparison between these two clips is not even close.



The MOS results confirm the blockiness results – MPEG-2 is clearly better, although inconsistent.



CLIP 5 *Geography*

Excerpt

Yale Repertory Theatre, New Haven, Connecticut, 4 November 1997

Conceived, Choreographed and Directed by Ralph Lemon

Videotaped by Johannes Holub Videographers

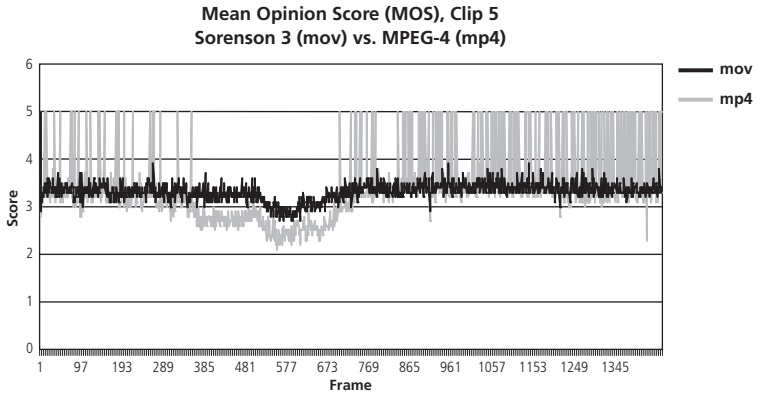
Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

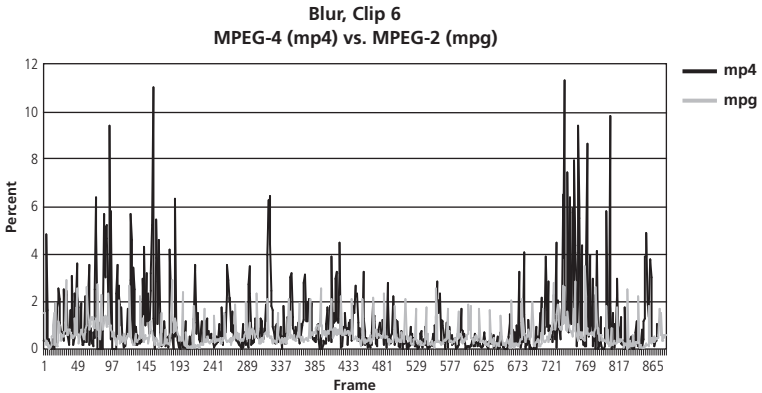
Why we were interested in this clip:

This piece has a very complex, intricate, and layered set. Overall the piece presents a high level of contrast between the performers and the space in which they are performing, and the camera does a pretty good job capturing the performance—but only on the close-up.

The codecs begin with similar colorfulness and with slight variations in the clip at the point where the mattress springs come into the frame. MPEG-4 becomes supersaturated, while Sorenson does an okay job. Toward the end of the clip the perceived colorfulness for Sorenson becomes supersaturated while MPEG-4 becomes less so.



MPEG-4 performs better than Sorenson, even though both have virtually the same bit rate. This graph shows that differences in codecs at the same bit rate can have substantial differences in perceived quality overall, even when single aspects such as colorfulness are virtually identical.



CLIP 6 *Oleg Tambulilingan or Bumblebee Dance*

Excerpt from The Dancers and Musicians of Bali

Town Hall, New York City, 22 March 1996

Danced by Cok Ratih Iriani and Made Lila Arsana

Videotaped by Johannes Holub Videographers

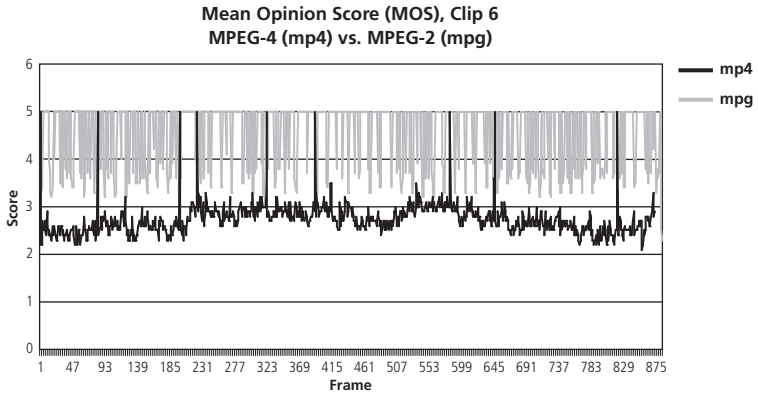
Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

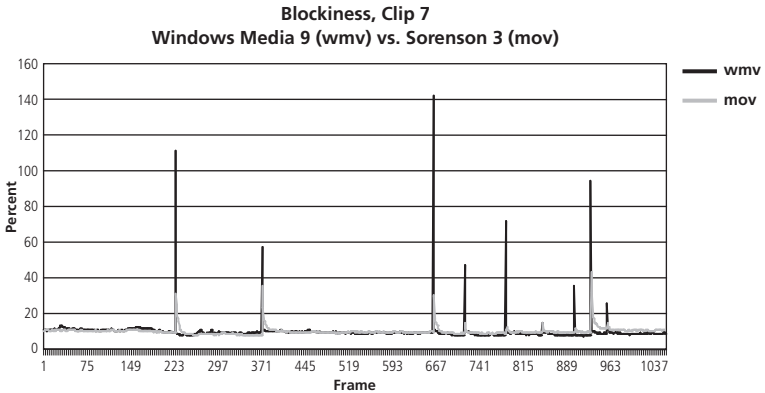
The dancer's outfit was so shiny and complex, we could not resist the desire to evaluate the artifacting caused by digital compression. We wanted to see how the main subject of this piece would fare compared to her more stationary onstage companions. The musicians are not moving around much, but their costumes are as detailed as hers.

This clip contains a single Balinese dancer with a very shiny, elaborate costume. Certain details could be lost in blur, for example the fine motion of the dancer's hands. This indicates the need for a high bit-rate codec to capture it all with as little loss as possible. The data suggests that 20Mbit MPEG-2 will do much better than a lower bit-rate MPEG-4. Obviously, in



most cases, a higher bit rate will produce a better result. However, even the high bit-rate MPEG-2 suffers from some blur, though not nearly as severely as the MPEG-4.

MOS scores for MPEG-2 are significantly higher for this clip, as might be expected, although the continual oscillation is of concern.



CLIP 7 *Improvisation*

Excerpt from Tap City New York City Tap Festival 2001

New 42nd Street Theater, 12 July 2001

Danced by Savion Glover and Gregory Hines

Videotaped by Charlie Steiner of Vagabond Video

Excerpt copied from Betacam SP

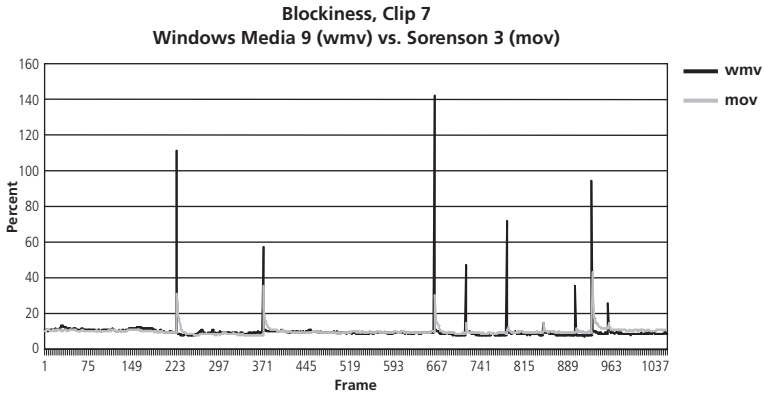
Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

There are a pair of dancers (Savion Glover and Gregory Hines), who dance in a space that is not well lit. We have some fast foot motion that could easily get blurry, as well as a multitone gradient background that could easily get very blocky.

Also, we were interested to see how well the hard, angled lines of the multi-planed stage area would hold up under compression: would they become jagged or would they remain smooth?

To compress the footage, Windows Media 9 relies on one frame being similar to the next. When there is a cut to a new frame with totally new information, the footage becomes predictably very blocky until the next full frame. This is evident in the spikes in the graph, which map exactly to the



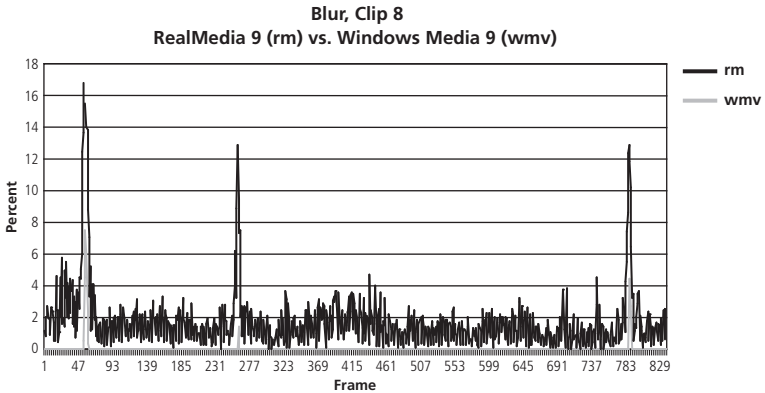
cuts in the footage. According to the data, the Sorenson 3 codec is doing a better job at looking ahead in the footage and predicting where it needs to process full frames.

The lit gradient background is blocky in both codecs, but appears to be much more pronounced in the MPEG-4 file.

In addition, the occasional flashes from cameras belonging to people in the audience make this scene more difficult for the Sorenson 3 and Windows Media to handle. The flashes are causing the entire background color to change, creating a very brief shadow of the dancers on the background. This changes the entire frame enough that it's difficult for either codec, but especially MPEG-4, to compress the file well.

Even though humans perceive the scenes as belonging to a coherent whole, the computer will see nothing similar.

MOS scores for this piece show that both results are similar, with the Sorenson scores being consistently better. Whether this difference is visually perceptible is questionable; the spikes at transition are more of a concern. Notice the difference for the same clip between this graph and the blockiness. Clearly blockiness is only one visually perceptible parameter, when weighted with other factors.



CLIP 8 *Primo Ballerino Stickman*

Excerpt from Deaths and Entrances

Mother, New York City, 4 November 1998

Martha@Mother with Richard Move

Created and Performed by Basil Twist

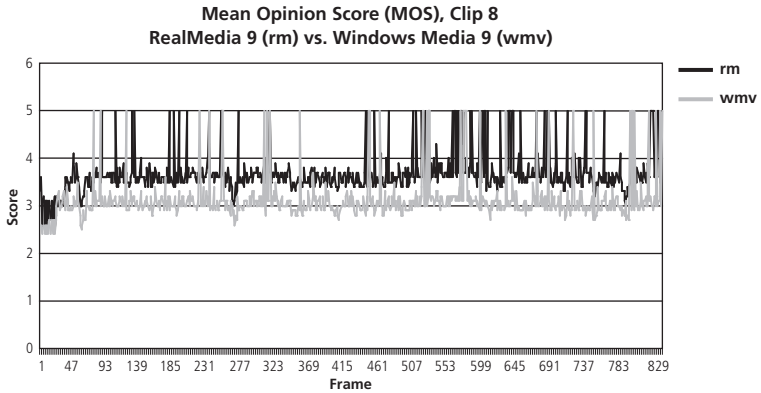
Videotaped by Charlie Steiner of Vagabond Video

Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

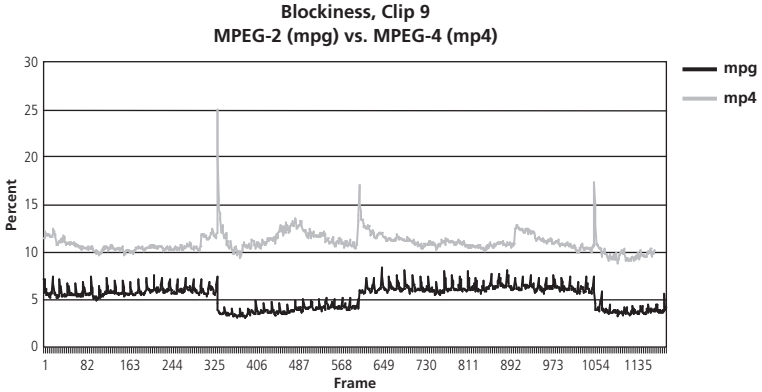
Why we were interested in this clip:

The “performer” in this piece is a puppet, powered by famous puppeteer Basil Twist. In addition to being in very high contrast to the background, the puppet performer is being held up by very thin strings. We wanted to know how well the strings would hold up under compression. Would they remain or would they disappear into the background? Would the motion of the puppet (who is very, very thin and fragile looking) maintain its delicacy or would it turn into a blocky mess?



According to the data, all spikes in blurriness correspond to pans and zooms of the camera, which while on a tripod does not have totally clean motion. RealMedia 9, in particular, blurs footage much more than Windows Media 9 as the camera moves.

Both codecs had similar results in terms of how they handled sharp transitions, which is not smooth. RealMedia does appear to outperform Windows Media, but the overall quality and spikes show a very similar viewing experience.

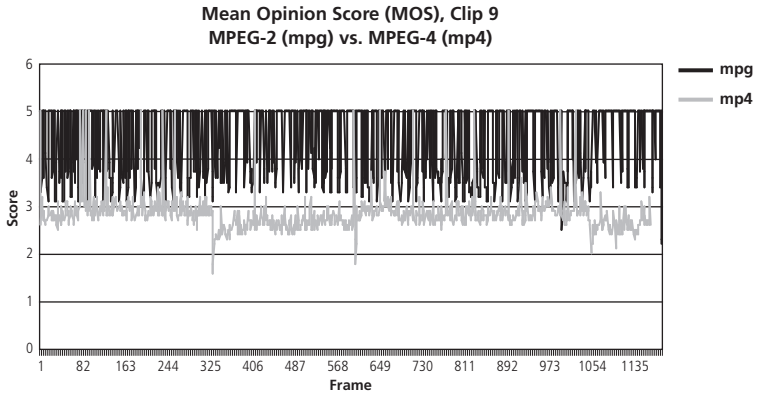


CLIP 9 *Inkblot*

Excerpt from Complexions—A Concept in Dance
Brooklyn Academy of Music Majestic Theater, 19 September 1997
Choreography by Dwight Rhoden
Artistic Direction by Dwight Rhoden and Desmond Richardson
Videotaped by Johannes Holub Videographers
Excerpt copied from Betacam SP
Courtesy of Dance Division, The New York Public Library for the
Performing Arts

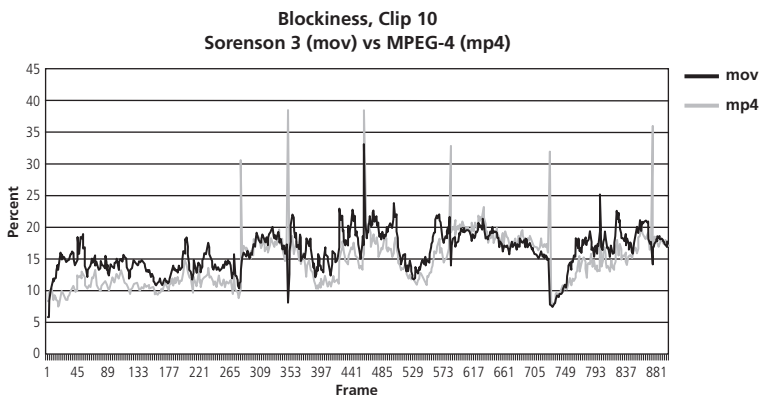
Why we were interested in this clip:

Large stage setting. Well-known choreographer. Definitely lit for stage and not camera. Fast motion with a large number of dancers performing different actions. Costumes are all single colors, but they are shiny and represent a wide variety of tones. Our interest was much more general on this clip—not so specific. It would be interesting to look at all the compressed versions of this clip to see where things broke. We could not anticipate anything specific at the time we picked the clip, but we knew it would look very poor when compressed.



With camera changes, the higher bit-rate MPEG-2 does not suffer from the same amount of blockiness as MPEG-4. In addition, the close-up camera (appearing second in the clip) is more effectively compressed by MPEG-2. This is evident from the drop in blockiness for MPEG-2 after frame 334, which is where the cameras switch. At this switch, MPEG-4 spikes sharply, indicating increased blockiness.

In this clip, MPEG-2 provides superior results, although MPEG-4 results are far more consistent from a perceived quality perspective.



CLIP 10 *Estampas y Tradiciones*

Excerpt from *Mano A Mano*, *Cultura Mexicana Sin Fronteras*
Haft Auditorium, Fashion Institute of Technology, New York City, 16
December 2001

Director, Francisco Nevarez Burgueno

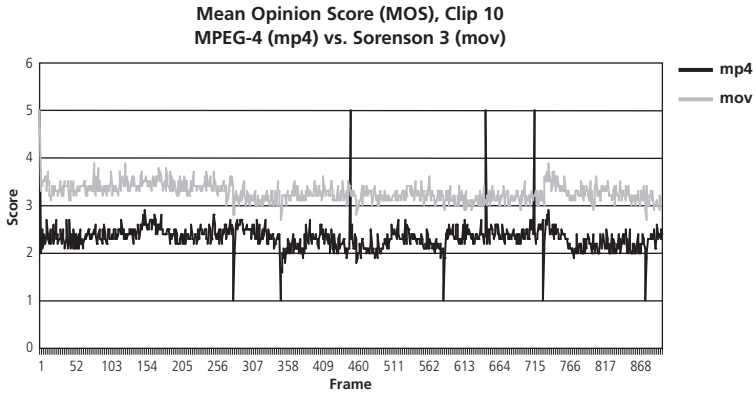
Videotaped by Francois Bernadi

Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

Fast motion combined with a swirl of complex costuming captured at two camera angles makes for a very exciting performance—in the theater. Unfortunately for these performers, one camera exposure is much better than the other. We were interested to see if one camera's footage would hold up better than the other.

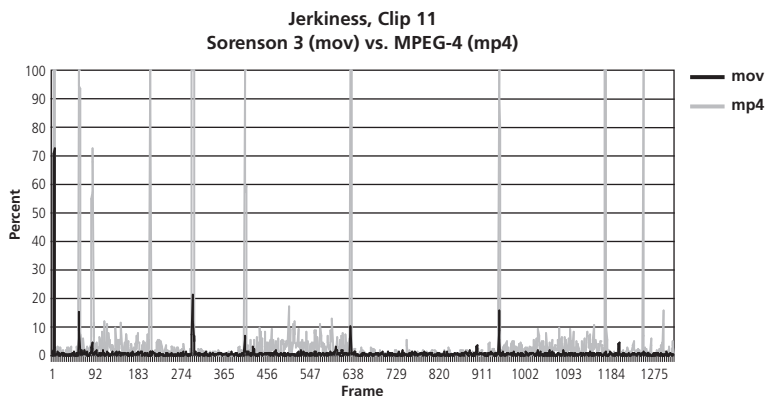


To compress the footage, MPEG-4 relies on one frame being similar to the next. When there is a cut to a new frame with totally new information, the footage will become predictably very blocky until the next full frame. This is evident in the spikes in the graph especially in MPEG-4, which map exactly to the cuts in the footage.

According to the data, the Sorenson 3 codec is doing a better job at looking ahead in the footage and predicting where it needs to process full frames.

The two cameras have different lighting exposures, which is making the job of both MPEG-4 and Sorenson even more difficult. Even though humans perceive the scenes as belonging to a coherent whole, the computer will see nothing similar.

Both codecs provide results that are consistent and tightly grouped, with only a few spikes. This is in contrast to many of the other results, in which the perceived quality oscillated significantly. Sorenson results are clearly better.



CLIP 11 *Bendiyan*

Thanksgiving dance, originally of the Ibalois tribe of Benguet
Excerpt from Pagbubunyi: A Celebration of Filipino Culture and
Heritage

Washington Irving High School, New York City, 2 April 2000

Performed by Bibak

Artistic Director, Erwin Kilip

Videotaped by Charlie Steiner of Vagabond Video

Excerpt copied from Betacam SP

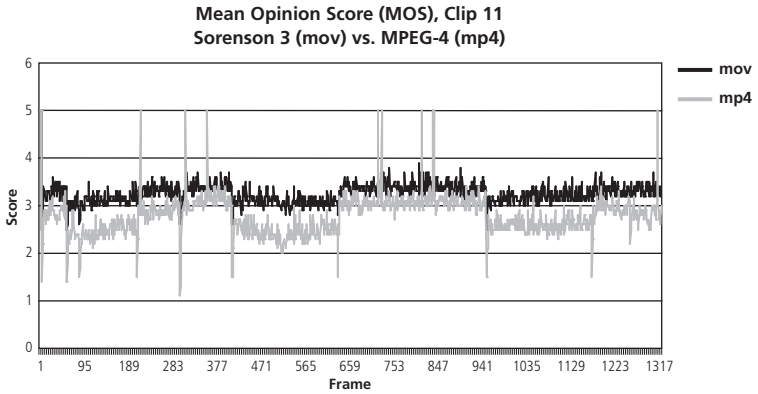
Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

Lots of people are going in different directions in an orderly fashion. Lots of skin or similar tone. It's bound to be blocky and jerky! Also, all costumes have horizontal lines in them and the dancers move in such a way that the lines all move together.

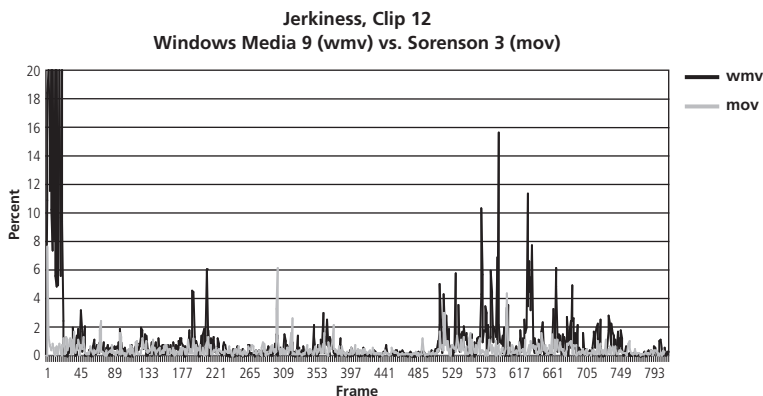
Both codecs have problems with jerkiness at the same moments in the clip. The data show, however, that the Sorenson codec is doing a much better job. Jerkiness in this footage corresponds to the cuts as well as to the moments of flash photography during the performance.

Analysis of the Tests Run on Footage



Overall, extreme blockiness in this footage contributes to the jerkiness.

Very similar results for the two different codecs.



CLIP 12 *While Going Forward*

Excerpt

A. J. Fletcher Opera Theater, Raleigh, North Carolina, 19 May 2001

Choreography by Tyler Walters

Carolina Ballet

Artistic Director, Robert Weiss

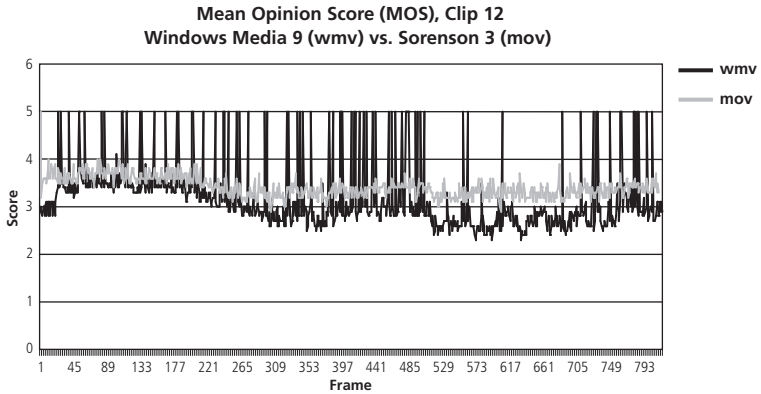
Videotaped by Warren Gentry & Associates, Inc.

Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

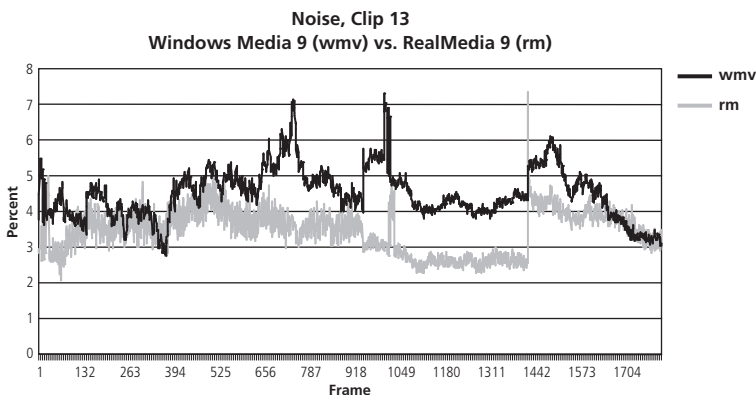
Why we were interested in this clip:

Here we have two dancers performing the same motions, side by side in costumes of contrasting colors. It is so dark that the light from the orchestra pit seems to seep in, in an obtrusive way. We are looking for blockiness in the costuming and blurring along the lines of the body and background. The stage itself is rather shiny. While “shiny stage” is not a prerequisite for performance of this piece, the effect on video is striking. We wondered if reflections of the dancers would show up at all.



The data indicate that Windows Media 9 has extreme difficulty with jerkiness during the first second or so of the clip. This could be due to the very high contrast of the scene. For the rest of the clip, however, Windows Media 9 continues to be outperformed by Sorenson 3. There is a distinct increase in perceptible blockiness around frame 570, when the dancers rise up abruptly after a short pause.

Once again, the oscillating nature of visually-perceived quality on Windows Media 9 is in stark contrast to the Sorenson 3 codec.



CLIP 13 *Abandoning Hope*

Excerpt from Triage

The Duke on 42nd Street, New York City, 17 March 2001

Created by Amy Sue Rosen and Derek Bernstein

Danced by Sally Bomer, Victoria Boomsma, Thom Fogarty, Sam Keany,
and Phillip Karg

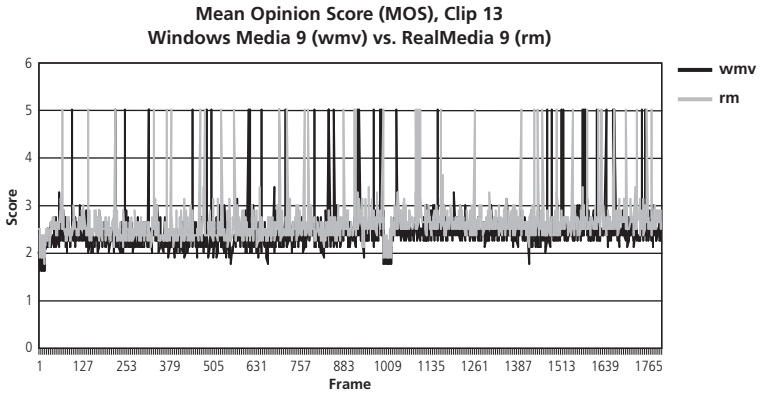
Videotaped by Charlie Steiner of Vagabond Video

Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

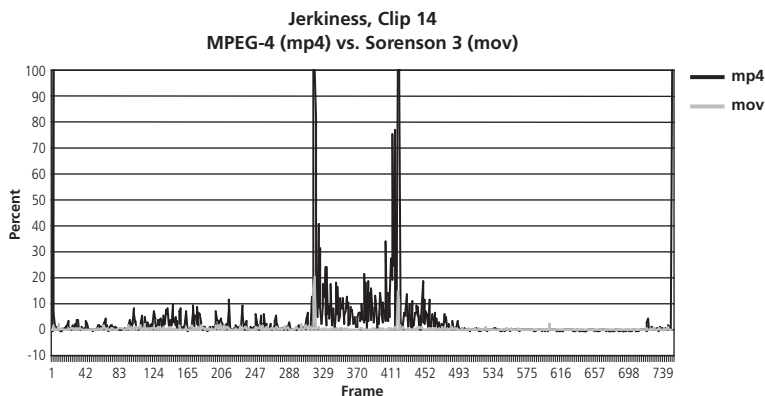
In this very morbid work, created by a woman who was dying of cancer, our primary interest was the mood-setting rain that is falling at the foot of the stage during the entire piece. We were curious to know how much compression it would take to make the rain look not as it was intended—or to make it disappear completely. Also of interest was to see how the gradient lighting at the foot of the stage compares to the stark darkness of the back of the stage. Look for blockiness up front. Finally, we were curious to see how the light faces of the dancers would fare against the stark black background—would they keep their detail?



Both RealMedia 9 and Windows Media 9 are introducing a fair amount of noise into the footage.

The rain at the foot of the stage (in front of the dancers) presents moments of brightness as light reflects on it, presenting challenges to both codecs. Blockiness in both codecs can be interpreted as noise, especially along the edges of the raindrops and the edges of the dancers' bodies.

Extremely similar results for both codecs in almost all aspects.



CLIP 14 *Nascimento*

Excerpt from Dance Women/Living Legends

Aaron Davis Hall, City College, New York City, 15 November 1997

Choreography by David Parsons

Dallas Black Dance Theatre

Founder and Artistic Director, Ann Williams

Videotaped by Robert Shepard

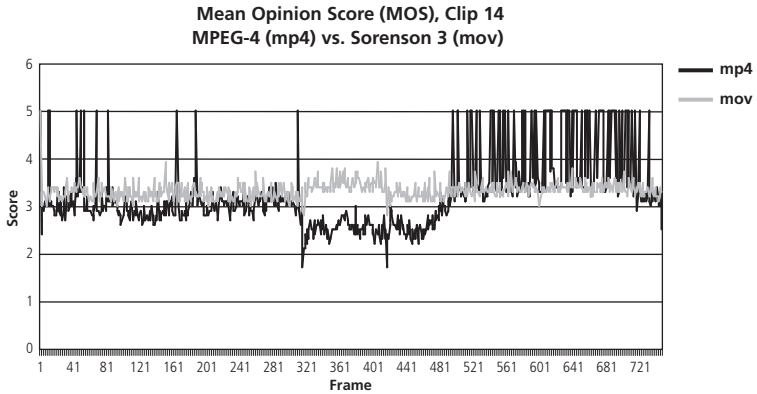
Excerpt copied from Betacam SP

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

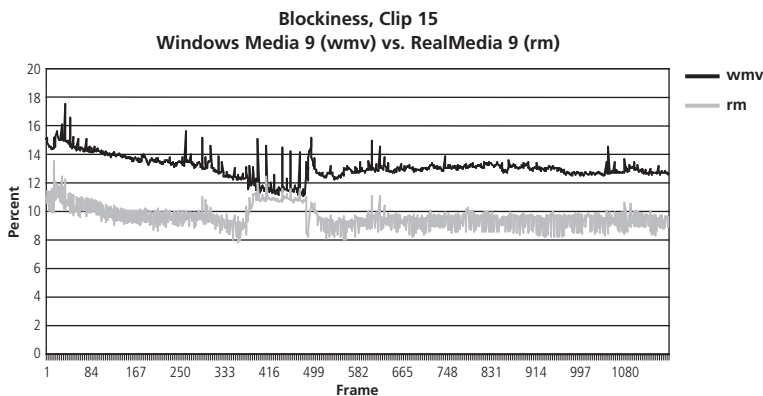
Why we were interested in this clip:

A well-known choreographer. This piece has a gradient background as well as multiple dancers.

Sorenson 3 is dealing much better with jerkiness in this clip. During the camera change to close-up, MPEG-4 is noticeably jerkier from the high motion of the dancer who fills the frame. Once the camera switches back, the Sorenson codec still performs quite well, while the additional dancers who enter the frame cause MPEG-4 to become perceptively more jerky.



This clip is a good example of visual inconsistency during a piece. The eye is drawn to this type of aggregate inconsistency in overall quality level. It is one thing to have rapid oscillation, but in this case there is pretty stable performance, which is interrupted by some extreme oscillation in MPEG-4. This shows how it is virtually impossible to predict codec performance, even within individual short pieces.



CLIP 15 *Not So Fast, Kid!*

Excerpt from Show Me

The Kitchen, New York City, 11 January 2001

Conceived and Choreographed by Cathy Weis

Cathy Weis Projects, Nova Productions from Skopje, Macedonia

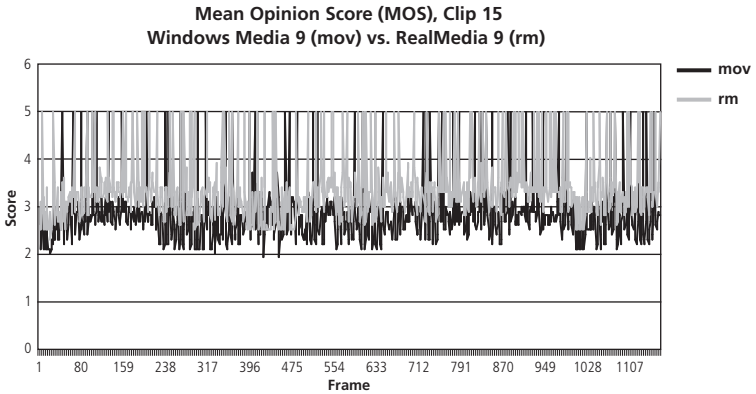
Videotaped by Charlie Steiner of Vagabond Video

Excerpt copied from DVCAM

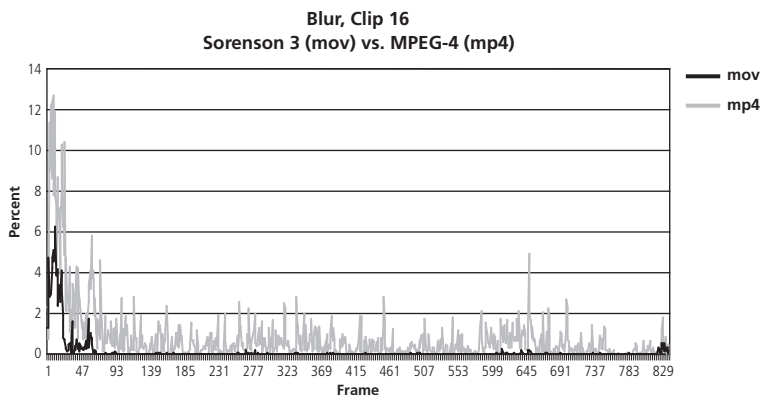
Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

This is truly a “multi-media” presentation. This piece combines live performers, performers at a remote location visible via a Webcam projected on a screen, and large cartoon drawings in the sets, as well as some projected text onstage. Each of these elements creates its own individual challenges to digital compression—but combined, the challenge is even greater. Be on the lookout for artifacts in certain areas of the frame, and different artifacts in other parts of the frame. In the first few seconds of the clip, the webcam projection shows its own blockiness, which is interpreted with the analysis software as general perceived blockiness.



Similar results from both codecs include wide oscillations of image quality. Blocky quality in both clearly hurts the perceived quality of the piece.



CLIP 16 *Not-About-AIDS-Dance*

Excerpt

The Kitchen, New York City, 15 December 1994

Choreography and Text by Neil Greenberg

Performed by Ellen Barnaby, Christopher Batenhorst, Neil Greenberg,

Justine Lynch, and Jo McKendry

Videotaped by Steve Brown of High Risk Productions

Excerpt copied from 3/4" Umatic

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

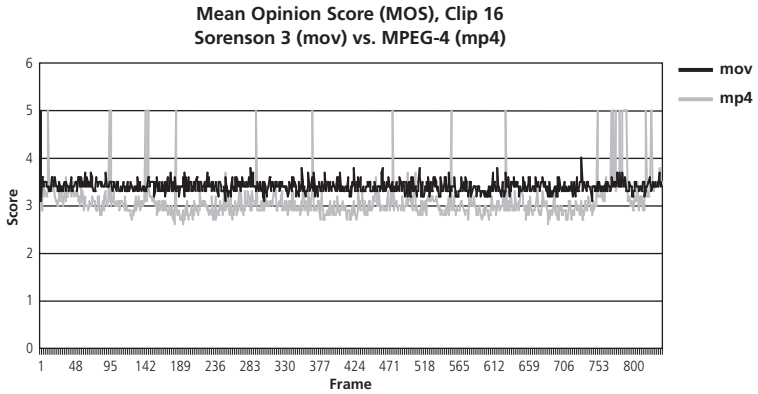
Why we were interested in this clip:

High contrast lighting a group of dancers dressed in white. Up lights in the back will become blocky.

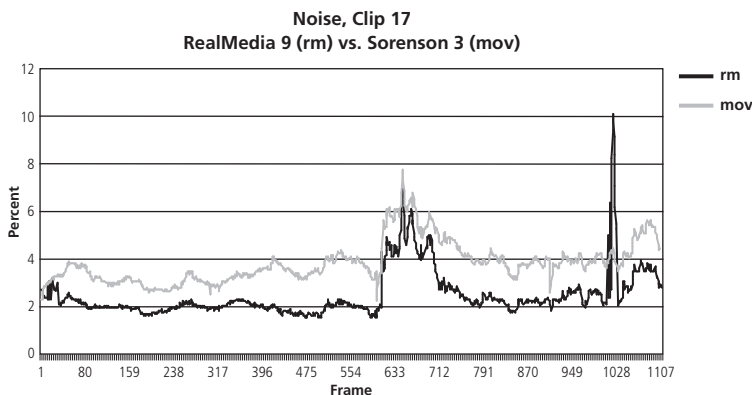
The initial camera pan in the first 50 or so frames of this clip produces marked blurriness in both Sorenson 3 and MPEG-4.

Careful viewing of this section revealed blurriness particularly in the background: the bricks of the theater wall illuminated harshly by spotlights.

Overall, however, Sorenson out-performs MPEG-4 in the ability to prevent motion from becoming blurry.



These results correlate fairly well to the blurry results noted above. Both systems encoded well, with tight quality grouping and quality that is very similar.



CLIP 17 *Menuet À Quatre*

Excerpt from Soirée Baroque en Haïti

Florence Gould Hall, New York City, 2 November 2003

Period Choreography by Catherine Turocy

New York Baroque Dance Company

Artistic Director, Catherine Turocy

Videotaped by Johannes Holub Videographers

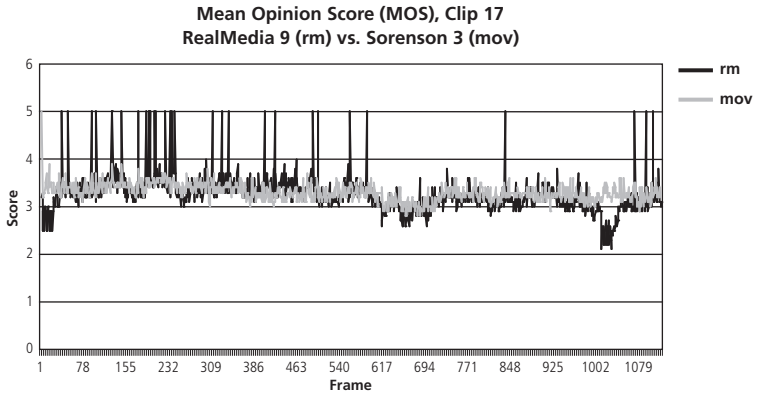
Excerpt copied from DVCAM

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

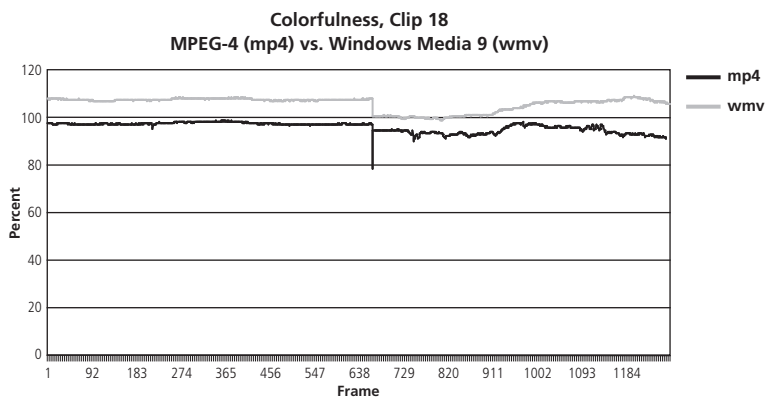
Why we were interested in this clip:

This presents a variety of skin tones and costuming. Also a group of dancers in a circle with attractive costumes. There are lots of hot spots in terms of lighting on the stage that will cause blockiness as well as a gradient. Complex patterns on dresses as well as the expressions on the faces of the dancers—keys to this genteel dance form—may be lost in compression.

Both RealMedia 9 and Sorenson 3 introduce a fair amount of noise into the footage. The noise becomes more pronounced as the camera zooms in slightly, which fills the frame more completely with the dancers. As the camera zooms back out slightly, there is another spike in noise.



RealMedia has some problems with the camera zooms in this clip, Sorenson handles them nicely.



CLIP 18 *Vodun Zépaule*

Excerpt from Soirée Baroque en Haïti

Florence Gould Hall, New York City, 2 November 2003

Choreography by Marcea Daiter

Dallas Black Dance Theatre

Founder and Artistic Director, Ann Williams

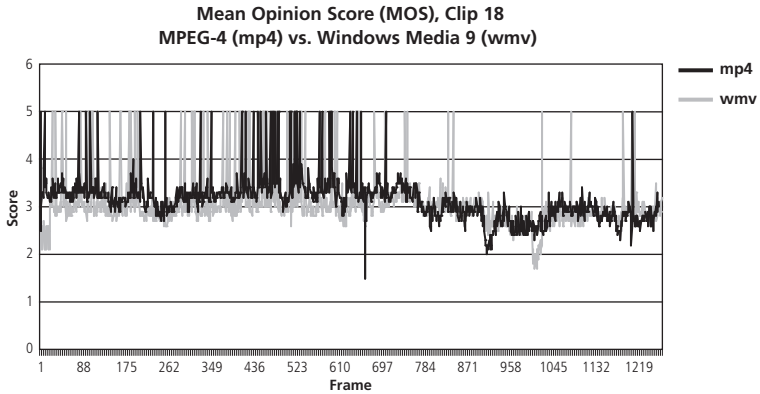
Videotaped by Johannes Holub Videographers

Excerpt copied from DVCAM

Courtesy of Dance Division, The New York Public Library for the
Performing Arts

Why we were interested in this clip:

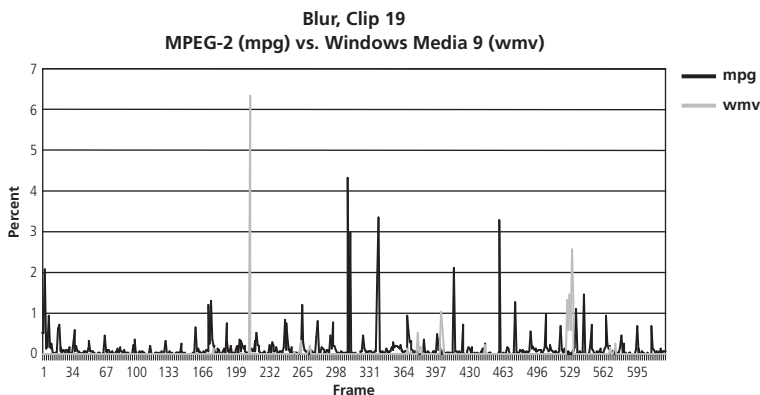
A key moment in the narrative of this piece is when the Trickster character blows magic dust on the two other dancers. We were interested to see if these crucial, detailed moments in the work could at all be preserved in compression. There is a gradient background that will get blocky. We also watched the gentle folds in the woman's dress and the man's pants for blockiness and stair stepping on the edges. We assumed that the mood-setting lighting pattern on the floor would quickly become blurry and blocky—detracting from the performers.



Both Windows Media 9 and MPEG-4 retain most of the original perceived colorfulness in the clip.

Windows Media 9 shows a higher degree of saturation than was actually in the original. The higher value expressed in the graph should be interpreted as loss of information, rather than value added.

These results are in marked contrast for those for colorfulness. Clearly, colorfulness is a low-weighted factor in the perception of overall quality. Both codecs provide similar results—although in this case the overall encoded quality is fairly tight in some sections with only a few spikes. Overall, this is unlikely to be a high-quality viewing experience.



CLIP 19 1992 Gala

Ted Shawn Theatre Presentation

Excerpt copied from Hi-8

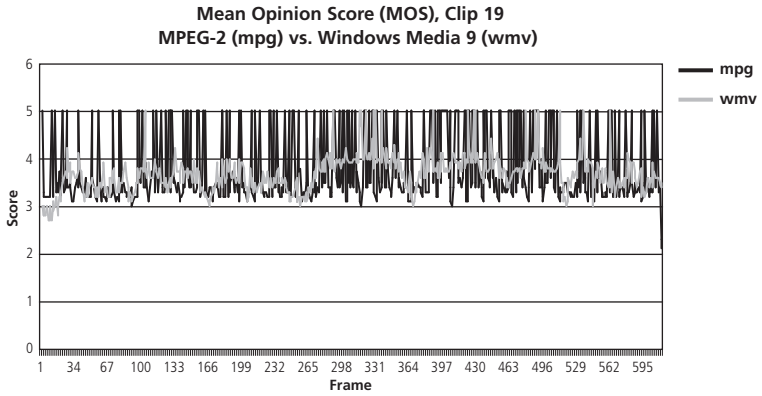
Courtesy of Jacob's Pillow Dance Festival

Why we were interested in this clip:

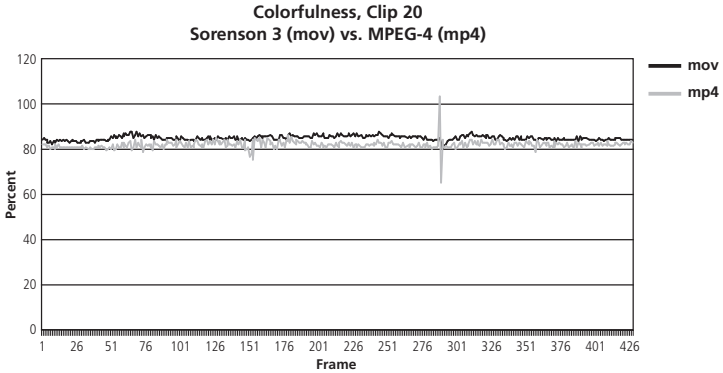
Any details in the performer's dress will most likely just disappear. Also, facial expressions will be much harder to discern. Mainly, however, we were interested to see if Hi-8 would at all hold up under compression. Jacob's Pillow—and presumably many other small archives—has Hi-8 and VHS. Jacob's Pillow does not have “professional” videotape formats.

This clip contains information almost exclusively at the end of the luminance scale. Its extremely high contrast footage is already at such low detail, from overexposure, that there are not many details available to be perceived as blurry.

Overall, both of these codecs exhibit low blur on these clips. However, there is blur associated with camera movement in both codecs.



In this clip, MPEG-2 does a better overall job with quality, but the results are very inconsistent.



CLIP 20 Chore

Student Showing, 25 June 1992

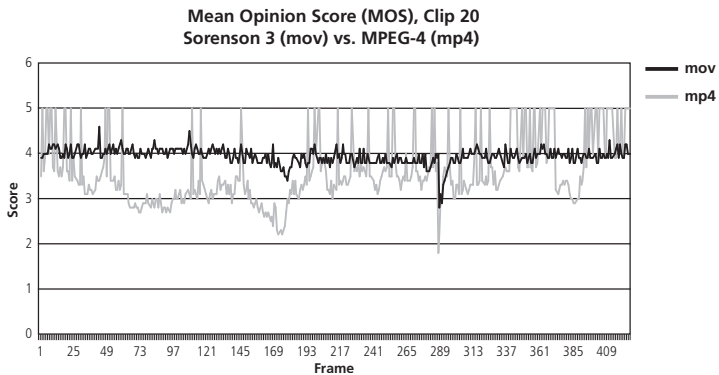
Excerpt copied from Hi-8

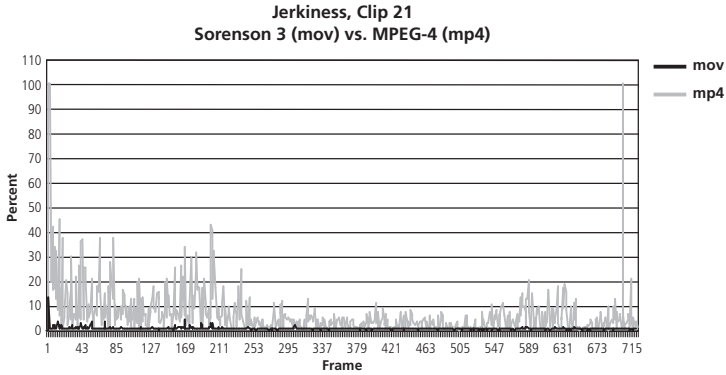
Courtesy of Jacob's Pillow Dance Festival

Why we were interested in this clip:

This Jacob's Pillow performance space presents some lighting challenges, as we see from the light coming from the side. The graph shows that both codecs performed almost identically in terms of dealing with color saturation. Sorenson 3 did a bit better at the key moment, as evident by the spike.

There were widely different results for the two codecs. MPEG-4 had a great deal of trouble with this clip.





CLIP 21 *Informance*

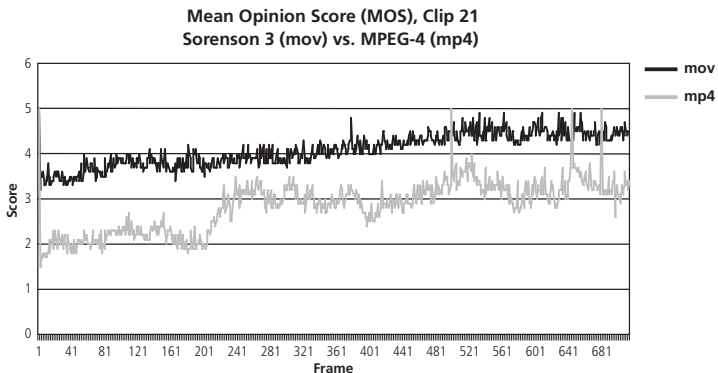
Choreography by Trisha Brown
Excerpt copied from VHS
Courtesy of Jacob's Pillow Dance Festival

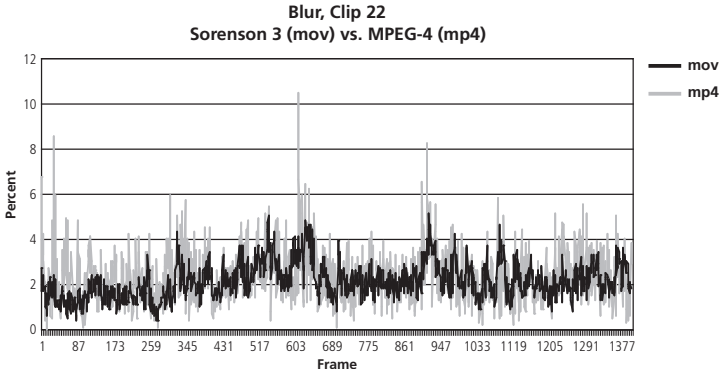
Why we were interested in this clip:

We wondered how well details survive compression when the originals have high contrast.

This clip demonstrates the superiority of the Sorenson 3 codec to MPEG-4 in dealing with perceived jerkiness. The multiple dancers do not seem to phase Sorenson 3, but MPEG-4 seems to be having a much more difficult time.

Sorenson is the clear winner on this clip.





CLIP 22 *Halau Hula O Hoakalei Ka Pa Hula Hawai'i Hula*

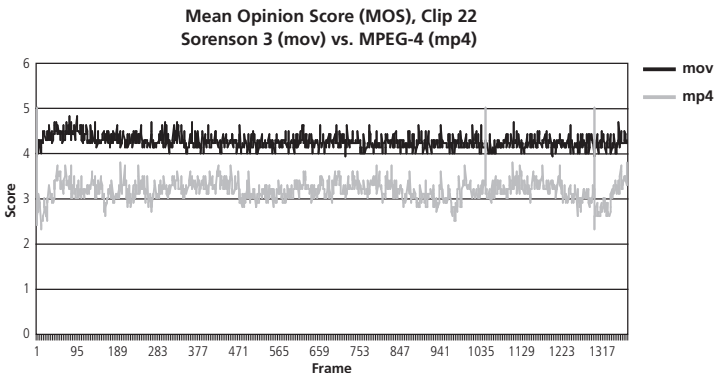
Excerpt from performance 3 August 1989 and workshop 4 August 1989
Excerpt copied from VHS
Courtesy of Jacob's Pillow Dance Festival

Why we were interested in this clip:

We wanted to see how well the details survive compression when the originals have high contrast.

In this clip, both codecs perform in a similar fashion. In general, Sorenson 3 outperforms MPEG-4; however, it can be seen that in some frames MPEG-4 is perceptibly less blurry.

Sorenson produced superior results.



SUMMARY ANALYSIS AND RECOMMENDATION

A chief goal of this report was to endorse a specific file format and codec to use for the preservation of dance material. Regarding file format, the Material Exchange Format (MXF) container format is recommended. Its focus on end-users—as opposed to broadcast organizations—and its requirement to contain digital media essence as well as its ability to contain metadata makes MXF the best choice to digitally preserve dance footage and ancillary information. This file format is further enhanced by being codec-agnostic, allowing for the use of any codec by which to encode and distribute dance materials.

After an exhaustive analysis, it became clear that there was no single lossy compressed solution that, consistently, was visually acceptable. We also determined that the criteria for preservation are significantly more rigorous than consumer-grade media or web content delivery, and none of the lossy compressed formats came close to performing the way we believe is required for this application. For this reason, we turned to lossless compression as the only viable option.

During the course of our study, JPEG2000 began emerging as a viable option for several reasons. JPEG2000 does offer the ability to do lossless compression. We tested this to make sure that the lossless compression was, in fact, mathematically lossless compression. In the past, the video industry has called lossy compression schemes “lossless,” which, while acceptable for the marketing purposes of the companies involved, are not factual. We were very pleased to find that after going through the JPEG2000 compression process, our .avi files were identical, when tested by the Genista software suite. For this reason alone, JPEG2000 was the only candidate format that met our criteria for mathematically lossless performance for archival purposes.

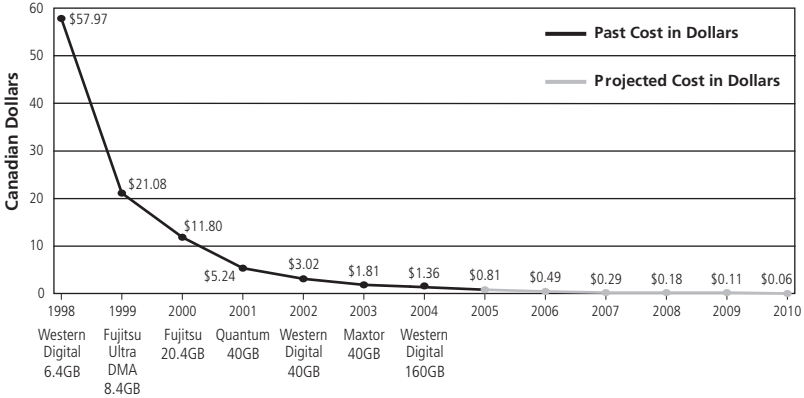
An additional benefit to JPEG2000 is that it is scaleable. This means that one can use the same “mother” lossless compressed file to create other lower quality files—which, while not acceptable for preservation, are very good candidates for distribution. So, from a technical point of view, JPEG2000 offers a good and viable solution for both preservation and access purposes. This is a *first* and it offers an extremely exciting option for both the dance community and for the larger archival community.

There are two major technical issues, however, that are real-world obstacles to the adoption of JPEG2000: (1) the cost of storage and (2) the availability of inexpensive real-time hardware for JPEG2000 codecs. We believe that both of these issues are currently being addressed in the marketplace.

It is beyond the scope of this report to do an extensive trend analysis of the cost of computer storage, particularly for the cost of hard disk storage. Nevertheless, a discussion of this subject is extremely pertinent to the problem at hand. Mathematically lossless compression, while it performs an essentially perfect job from a file preservation point of view, is less efficient than other approaches, since it has a compression ratio of approximately 3:1. Further, experts have been working on lossless compression algorithms for quite some time, because of their use in the larger information technology (IT) environment, and while breakthroughs are always possible, it is unlikely that a breakthrough will occur that gives lossless compression the kind of ratio yields that lossy compression can easily generate. We therefore need to look elsewhere to determine whether there is another way to accomplish our preservation goals at a cost both realistic and affordable for the dance community.

We do not think that a revolution in lossless compression-yield ratios is likely. Nevertheless, we do believe that the constant and consistent trend in the reduction of the cost of hard drives will make for an economic change so significant that poorer yields will become much less meaningful.

The Declining Cost of Storage:
Past, Present, and Future



The graph shows the steeply decreasing cost of storage from 1998 to 2004, where the cost per gigabyte (GB) of storage decreased from about \$60Cdn to \$1.36Cdn. (Canadian dollars were used because we had real data from retail stores for specific drives from this period, which was unavailable in the U.S. marketplace). Perhaps even more relevant are our own observations during the period of our study: we found that raw disk storage cost (the cost of an unformatted hard drive in gigabytes) decreased from \$1 per gigabyte (U.S.) in November 2003 to \$.79 per gigabyte in May of 2004—a period of only six months.

We believe that it is fair and reasonable to count on the continuing trend of decrease in cost per gigabyte, based on current trends. Therefore, we can look at the cost of storage through a very short telescope (six years) to try to forecast the approximate cost of using mathematically lossless compression to archive video material. Based on our forecast in the graph above, we think that it is likely that the cost will be in the area of approximately \$.06 per gigabyte. If we are off even by 100% the cost will be only \$.12. There is great industry support in the literature for this forecast, and industry publications are basing the future growth of the industry on the continuing downward trend in costs of storage per gigabyte. There is no shortage of industry speculation in this particular area, where, for example, in the Febru-

ary 2004 issue of *PC Magazine*, a prediction is made of 700GB as the normal configuration for personal computers (PCs) in 2007. The recent introduction, in March of 2004, of a 400GB single drive by Hitachi (formerly the highest capacity drives readily available in an inexpensive format were 300GB), further supports the continuing evolution of increasing storage quantities with the simultaneous reduction in cost.

While video contains a great deal of information, it is well defined, and as data capacity continues to expand with decreasing cost, we can forecast a time in the near future when storage cost as an element of overall cost is no longer very significant. Currently, about 1 hour of content can be mathematically losslessly compressed into approximately 25 gigabytes of space. That is a large file, and from a cost point of view, today's raw storage cost for that much data is \$19.75 (U.S.). A digital Betacam tape that stores a similar 1 hour of content costs over \$30. Videotape costs for professional formats have not made very significant recent cost decreases, and, in our opinion, are unlikely to. While there is, of course, a great deal of infrastructure involved in recording a file on hard drive, the same holds true with videotape. However, we believe that the huge quantities of hard drives being manufactured and the continual push of the industry will continue the trend that has been with us now for a very long time.

If our forecasts are close to accurate, by 2010 the cost of recording an hour of content will be well under \$2—a price that is affordable for the dance community. We therefore believe that this makes a very persuasive argument for the dance community—to anticipate and plan on decreasing storage cost as part of a preservation and distribution strategy for dance material.

Our other reservation was the current availability of inexpensive real-time JPEG2000 hardware encoders, to allow for the ready compression of the materials. In this area, too, we have reason to be very optimistic. The way to accomplish this task is for JPEG2000 to be available as hardware codecs. Recently, Analog Devices has announced and actually has begun delivery

of JPEG2000 hardware encoding and decoding chips. Mass production of chips to enable the ready and inexpensive incorporation of JPEG2000 compression in a wide variety of devices will insure availability. Extremely encouraging is the fact that JPEG2000 is an open standard; it thereby overcomes our concerns of obsolescence by providing a way to decode files over time combined with hardware to do it in real-time. During the last weeks of this study, Media Matters was able to evaluate a prototype device that, in fact, does encode and decode JPEG2000 at NTSC video rates in real time. Frankly, we were very impressed.

When we started Phase I of this process in 2002, we did not have a great deal of confidence that we would or could find a solution. The work that we did with lossy compression in many ways empirically verified what we then believed to be the case: that while fine for some distribution applications, lossy compression is wholly unsuitable for preservation purposes. This contention has been confirmed. What we did not anticipate was that a new industry standard would enable the archival community to rethink its direction and consider, seriously and perhaps for the first time, that there really was a viable alternative on the horizon: JPEG2000 lossless compression both satisfies the needs for preservation at the highest quality levels and is affordable enough to implement.

We have no guarantees that computer storage will continue to decrease in cost per gigabyte, but we deem it extremely probable. For this reason we would encourage the dance and other archival communities to plan a transition to losslessly compressed file storage, based on industry trends that, for many years, have continually delivered storage at decreasing prices. We find that the availability of an open standard is a very important step and that cost-effective hardware will allow for a preservation strategy that is affordable and implementable.

APPENDIX

ANALYTIC TOOL—GENISTA’S MEDIA OPTIMACY

For the experiments in Phase II of the Digital Video Preservation Reformatting Project, it was determined that “just watching video footage compressed via different methods to see what looks best” was not going to be enough. Tools were needed to examine the files on the signal level, in order to establish where and when in a file artifacts appear as the result of compression.

Along with a rise in new methods to deliver digital content via broadcast and streaming, new companies are developing that will examine the quality of the delivered files—when they are delivered. Companies are also developing tools that examine compressed video and audio and compare them, electronically, to the original, uncompressed footage.

One company is Genista, a young Tokyo-based company focused on creating accurate and easy-to-use software tools that measure the audible and visible artifacts caused by compression and transmission. Perceptual quality measurement tools, such as Genista’s Media Optimacy, have enabled content providers to develop associated network-delivery mechanisms for the best possible audience experience.

The following, excerpted from Genista’s Media Optimacy user manual, describes how the software works and how it draws the conclusions it draws.

VIDEO QUALITY METRICS

Genista has developed a set of metrics for measuring the quality of digital video and still images. Genista’s quality metrics measure the typical artifacts introduced by processing (notably compression) and transport of digi-

tal video. Additionally, a metric exists to make a prediction of Mean Opinion Score (MOS) (i.e., reproducing the results of human subjective tests on overall image quality).

Genista metrics are not merely based on network statistics or network performance parameters such as packet loss. Instead, they take into account the image content and frame data of the video resulting from the given coding and transmission conditions. The metrics can be divided into spatial and temporal metrics. Spatial metrics, such as blockiness, perform their measurements on a frame-by-frame basis, returning a result for each frame measured. Temporal metrics, such as jerkiness, look at two or more consecutive frames simultaneously to obtain a measurement. MOS prediction takes into account both spatial and temporal aspects.

RELATIVE AND ABSOLUTE METRICS

Video quality measures can be divided into relative (full-reference, FR) metrics and absolute (non-reference, NR) metrics. FR metrics compare a compressed or otherwise processed video directly with the original, whereas NR metrics analyze any video without the need for a reference, using only the data contained in the clip under test.

Full-reference metrics are suitable for intrusive, out-of-service measurement of video quality. They provide video quality monitoring and management at locations where both the reference video and the processed video are available (e.g. at the encoder). They also lend themselves to applications such as encoder rate control.

Non-reference metrics target real-time measurement of streaming video. Such metrics enable the measurement of streaming video quality at any point in the content production and delivery chain. They are particularly useful for monitoring quality variations due to network problems, as well as for applications where service level agreements and quality control are

required. Another possible application is characterization of the reference content prior to encoding or processing. Currently non-reference metrics exist to measure jerkiness, blockiness, blur, and MOS.

The Metrics. The metrics provided by Genista comprise three categories:

Fidelity metrics measure the mathematical difference between processed and reference video.

Spatiotemporal metrics are defined by the ANSI standard (as discussed below).

Perceptual metrics include a prediction of MOS, which provides an overall perceptual quality in MOS scale.

Each of Genista’s metrics is described in more detail in the following sections.

Fidelity Metrics are widely used and represent arithmetic measures of the distance between processed and reference video. They are full-reference metrics by definition. Although fidelity metrics are very popular in the image- and video-processing world, they do not take into account human perception

Spatiotemporal Metrics rely on algorithms defined by recommendations from the American National Standards Institute (ANSI). Their recommendation represents an attempt by a standards body to define objective measures that serve as a basis for the measurement of video quality. These include the following:

Fidelity Metrics

Fidelity Metric Type Description

PSNR FR, spatial Peak Signal to Noise Ratio (luminance).

SNR FR, spatial Signal to Noise Ratio (luminance).

RMSE FR, spatial Root Mean Square Error (luminance).

Color PSNR FR, spatial PSNR from CIE ΔE_{ab} or ΔE_{94}

METRIC TYPE DESCRIPTION

Motion energy difference:

FR, temporal. Added motion energy indicates error blocks, noise.

Repeated frames FR, temporal. Lost motion energy indicates jerkiness.

Edge energy difference:

FR, spatial. Indicates dropped or repeated frames.

Horizontal and vertical edges:

FR, spatial. Added edge energy indicates edge noise, blockiness, and noise.

Spatial frequencies difference:

Lost edge energy indicates blur.

Perceptual Metrics. Genista’s perceptual quality metrics measure specific artifacts introduced into the video as perceived by a human viewer. These artifacts are well known, and are easily recognized even by nonexperts. The aim of these metrics is to provide an automatic measure of those artifacts that viewers will perceive, in a way that is correlated with human perception. Additionally, a metric exists to make a prediction of Mean Opinion Score (MOS), i.e., reproducing the results of human subjective tests.

Jerkiness is a perceptual measure of frozen pictures or motion that does not look smooth. The primary causes of jerkiness are network congestion and/or packet loss. It can also be introduced by the encoder dropping or repeating entire frames in an effort to achieve the given bit-rate constraints. A reduced frame rate can also create the perception of jerky video.

Lower levels of jerkiness can be perceived when subregions of the image appear to be moving in a jerky way. This can be caused by a variety of factors. For example, it can become apparent in smooth regions where changing contours or blocking artifacts can create the appearance of jerky motion.

Genista has developed both FR and NR jerkiness metrics.

Blockiness is a perceptual measure of the block structure that is common to all discrete cosine transform-based (DCT) image compression techniques. The DCT is typically performed on 8 x 8 blocks in the frame, and the coefficients in each block are quantized separately, leading to artificial horizontal and vertical borders between these blocks. Blockiness can also be caused by transmission errors, which often affect entire blocks in the video. Genista has developed both FR and NR blockiness metrics.

Blur is a perceptual measure of the loss of fine detail and the smearing of edges in the video. It is due to the attenuation of high frequencies at some stage of the recording or encoding process. It is one of the main artifacts of wavelet-based compression techniques, such as JPEG2000, where transmission errors or packet loss can also induce blur. DCT-based compression schemes are also affected by this artifact, albeit to a lesser extent (JPEG, MPEG). Other important sources of blur are low-pass filtering (e.g., analog VHS tape recording), out-of-focus cameras, or high motion (leading to motion blur). Genista has developed both FR and NR Blur metrics. Subjective experiments with images of different blur and JPEG2000-compressed images show a correlation of up to 96% between Genista’s blur metric and perceived blur.

Noise is a perceptual measure of high-frequency distortions in the form of spurious pixels. It is most noticeable in smooth regions and around edges (edge noise). This can arise from noisy recording equipment (analog tape recordings are usually quite noisy), the compression process, where certain types of image content introduce noise-like artifacts, or from transmission errors (especially uncorrected bit errors).

Ringling is a perceptual measure of ripples, typically seen around high-contrast edges in otherwise smooth regions (the technical cause for this is referred to as Gibb’s phenomenon). Ringing artifacts are very common in wavelet-based compression schemes (e.g., JPEG2000), but they also appear to a slightly lesser extent in DCT-based compression techniques (e.g., JPEG, MPEG).

Colorfulness. The colorfulness of an image describes the intensity or saturation of colors as well as the spread and distribution of individual colors in the image. The range and saturation of colors often suffer after compression. Subjective experiments with images of different colorfulness have shown a correlation of 93% between Genista’s colorfulness metric and perceived colorfulness.

Watermarking Artifacts. Digital watermarking of digital images and video content is becoming an increasingly important way for content producers and providers to protect their digital content without compromising the extent of its distribution. One of the most important factors when watermarking content is to minimize the perceptual impact of the watermark on the content. The ideal way to do this is to use perceptually based metrics that can reproduce the impact of the watermark on a human observer.

Based on five watermarking algorithms, Genista has developed metrics that offer perceptual measurements of two different artifact types present in digital watermarks:

- **Watermarking Flicker:** This measures visible temporal effects emerging from the relationship between successive frames of watermarked content. Such artifacts are particularly disturbing when video is watermarked with schemes optimized for still images. In such a scenario, the watermark changes between frames in a way that induces a very obvious “flicker” when a video is viewed. Genista’s watermarking flicker metric has been optimized using subjective test data taken from human observation of watermarked video, and has been confirmed to have a correlation of 95% with subjective data (compared to 54% for PSNR).

- **Watermarking Noise:** Since watermarking involves the manipulation of some fraction of the pixels in the digital content of an image, noise is a typical artifact produced by the procedure. Genista’s watermarking noise metric has been optimized for the type of noise typically induced by the addition in video content of a watermark. It has been optimized, using subjective test data taken from human observation of watermarked video, and has been confirmed to have a correlation of 81% with subjective data (compared to 41% for PSNR).

MOS Prediction. MOS is the Mean Opinion Score obtained from experiments with human subjects. Genista’s MOS predictions are metrics that correlate with human perception of video quality and thus with the output of subjective test results. Genista’s MOS prediction uses some of the above-mentioned perceptual metrics to construct a metric that represents the perceived quality of video content.

A set of subjective test data has been used to confirm the high correlation that this measure has with MOS from subjective tests. It should be noted that the accuracy with which this metric reproduces subjective MOS is necessarily dependent upon the type of content used. It has been demonstrated that for typical video content, covering a wide range of motion and texture ranges as well as common PC video codecs, the correlation of the metric with subjective MOS is significantly higher than PSNR.