NATIONAL ENDOWMENT FOR THE
Humanities

DIVISION OF PRESERVATION AND ACCESS

Narrative Section of a Successful Application

The attached document contains the grant narrative and selected portions of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Prospective applicants should consult the Preservation and Access Program application guidelines at http://www.neh.gov/grants/preservation/preservation-and-access-research-and-development for instructions. Applicants are also strongly encouraged to consult with the NEH Division of Preservation and Access Programs staff well before a grant deadline.

Note: The attachment only contains the grant narrative and selected portions, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: An Open Source Application for Image-Based Digital Reproduction of Optical Film Sound

Institution: University of South Carolina Research Foundation

Project Director: Gregory James Wilsbacher

Grant Program: Preservation and Access Research and Development

An Open Source Application for Image-Based Digital Reproduction of Optical Film Sound

A proposal to the National Endowment for the Humanities
Preservation and Access Division, Research and Development Awards

The University of South Carolina Moving Image Research Collections (MIRC) is collaborating with research faculty from the University's Interdisciplinary Mathematics Institute (IMI) to develop an open-source application to directly decode optical sound tracks from motion picture films to digital audio tracks that retain as much as possible of the original analog film quality. We seek $362,125 from the National Endowment for the Humanities over a three-year period to take our proven concept through development to deployment as a freely available application. The total budget for the project is $465,905.

When complete, the software will have two user modes: standard and custom. The standard mode will be designed for non-technical users. In this mode minimal interaction will be required and most of the processes will be completely automated. The custom mode will provide access to the parameters governing the implemented routines. It will require users to have a basic understanding of the underlying algorithms that will allow them to experiment in an intelligent way. Relevant explanations and guidance will be provided in an electronic manual.

SIGNIFICANCE

The twentieth century was the century of film. Preservation of the twentieth-century's cultural and historical record for future generations will have been fundamentally incomplete if that century's film production is not migrated into the digital realm in a manner that provides faithful surrogates of the celluloid record. The National Endowment for the Humanities, The Library of Congress, and other federal agencies have recognized this imperative and funded film preservation for the past three decades, but film-to-film preservation is getting more, not less expensive.[1]

The availability to scholars, students, and the general public of motion picture film has reached a crisis state. There is simply too much film, too little functional exhibition or screening equipment, and dwindling professional expertise on the maintenance of analog screening equipment. Even when equipment is available, aging film prints may no longer be suitable for use in machines designed for new film elements.

Digitizing film as a means of access and preservation has become an increasingly important alternative to traditional photochemical preservation. Procedures for converting optical sound tracks, however, have not kept pace with the development of image scanning techniques in part because technological advancements tend to be driven by the needs of the commercial film industry, not archives.

Commercial film production separates image and audio content into two distinct workflows married only at the time of the final product. Digital film scanners now in production are designed with this segregated workflow in mind (see Appendix A). As a result, digitally scanned sound film results in only the image

---

[1] The National Film Preservation Foundation (NFPF) is sole federal funding agency dedicated exclusively to film preservation. According to its most recent annual report, since its inception in 1998 NFPF has sponsored film preservation for over 200 organizations and preserved more than 750 films ("Report to the U. S. Congress for the Year Ending December 31, 2009). The National Endowment for the Humanities has also played a significant role in film preservation. The Newsfilm Library (now MIRC), for example, was awarded a Preservation grant in 2003 to make film preservation copies of over 200,000 feet of Fox newsreel material.

content of the celluloid record being converted into a digital surrogate. No sound information is converted to digital audio through a scanning technique.[2] When optical sound tracks are digitized, the standard procedure requires the tracks to be processed on aging analog equipment that is expensive to maintain and operate.

Unless cost-effective, widely available, and easy to use tools are created to aide in the conversion of optical film sound to digital film sound, large portions of the motion picture records important to humanistic study will not be available in a form suitable for researchers.

The MIRC/IMI optical sound conversion tool is designed to increase number of motion picture sound films migrated to digital sound film by radically decreasing the cost of this conversion. It accomplishes this by creating a one-pass scanning workflow that results in a synchronized digital sound film.

It is in the best interest of the nation if such a tool is developed by the non-profit sector and made freely available. If such a tool were protected by patents and constrained by licensing fees its impact on film preservation would be greatly diminished. MIRC/IMI are committed to providing such open access to the optical sound conversion tool created with the support of the National Endowment for the Humanities.

Background of the Optical Sound Preservation Problem

From 1900 to 1980, celluloid motion picture film was the principle means of conveying visual imagery in motion. Since the introduction of viable synchronized sound-film systems in the late 1920s (for 35mm film) and 1930s (for 16mm film), film sound has been a central component of moving image cultural records. Synchronized film sound was first achieved through the creation of optical sound records or separate audio discs. Although magnetic sound recording was developed in the 1930s and used extensively in the 1950s, especially in the case of wide screen formats, optical sound systems remained popular and dominated the 16mm film distribution market through to its demise.

Although a wide array of optical sound tracks were developed during this period, they were all based on the same principle of capturing on film the analog wave signal of natural sound. Variable density sound tracks saw the first widespread commercial success in 1928 with Fox Films', *Fox Movietone News* (figures 1 and 2). The patent and frequency limitations of variable density technology resulted in the dominance of variable area sound tracks (Figures 3 to 6).

[2] Lasergrapics' recently released scanner, "The Director," includes in its specifications an optional optical sound head. Requests to the company for further information about "The Director's" sound capabilities have not been answered at the time of the grant's submission.

Exemption 4

Fig. 1: Variable density track, *Fox Movietone News,* Bobby Jones in New York [1934] University of South Carolina.

Exemption 4

Fig, 2: Variable density track, *Fox Movietone News,* Bonus Army Marchers in Washington, DC [1932]. University of South Carolina.

Exemption 4

Fig. 3: Variable area track (unilateral). *C. E. Feltner Collection,* unidentified (ca. 1930). University of South Carolina

Exemption 4

Fig. 4: variable area track (bilateral, push pull). *C. E. Feltner Collection,* Dedication of Yankee Clipper [1929]. University of South Carolina

Exemption 4

Fig 5: Variable area (bilateral). *C. E. Feltner Collection,* Post-war Japan (ca. 1945). University of South Carolina.

Exemption 4

Rg. 6: variable area (single) on 16mm, *Local Television News Collection,* Dr. King Speaking in Kingstree, SC (1966). University of South Carolina.

Our sound moving image heritage captured with optical tracks includes published theatrical, educational, industrial, advertising, experimental and independent films. It also includes historically important unpublished newsreel and local television news material. Recently, Rick Prelinger, a noted expert on nontheatrical films, estimated that over 300,000 sponsored films were produced in the United States alone.[3] This category only comprises a portion of the nontheatrical market, leaving out--among other categories--the substantial legacy of 16mm television newsfilm produced by local affiliates throughout the country from the late 1940s through the early 1980s. In short, the amount of sound films produced in the past century far exceeds our capacity to reprint each film photochemically.

Even while film was the dominant mode of moving image distribution, televised broadcast of motion picture film required a system for converting the celluloid image and sound information running at 24 frames per second (fps) to NTSC video running at 29.97 fps. Film-to-video film chains and telecines were designed to meet this need. These machines were envisioned as post-distribution devices to extend motion picture film into emerging television and video markets. All telecines and most film chains converted film image and film sound to NTSC Standard Definition (SD) video in one pass and in real time (24 fps), providing a ready-to-view video signal.

Film to video transfer has been providing the general public with extended access to our collective film culture for thirty years--although only a small portion of the non-theatrical film library was made available in video. Standard definition (SD) video, however, has always lagged behind the image resolution of the film element and the conversion from 24 fps to 29.97 fps has made SD video an unacceptable surrogate for scholarly research and aesthetic fidelity. Accomplishing the conversion from 24 fps film to 29.97 fps, a telecine compensates for the frame rate difference by way of a 2:3 pulldown extension of the video fields. This procedure distorts the original film record by generating video fields not present in the original film (see figure 7).

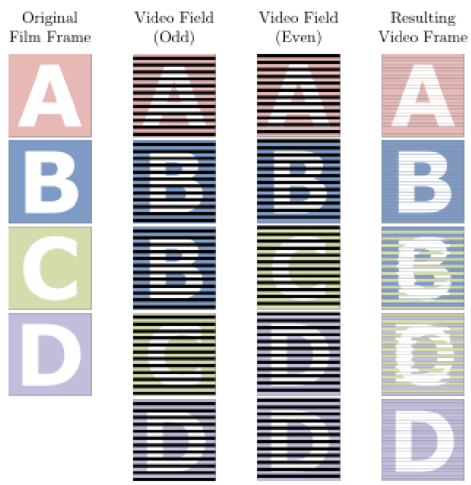


Figure 7. 2:3 telecine pulldown interpolates 24 fps film transfer to 29.97 fps video, producing a surrogate that does not accurately reflect the original film element. Illustration taken from "Telecine," *Wikipedia*, accessed June 11, 2010.

Although the transition to HD video has brought the video image closer to film, it still falls short of film's resolving capacity and continues the problem of frame rate interpolation.[4] As long as celluloid film projection was commonly available, video could be accepted as an imperfect but valuable *component* of an access strategy. The loss of routine access to correctly projected film, however, highlights the need to find a replacement for SD and HD video as a means of access to our film heritage.

[3] Prelinger, Rick. *The Field Guide to Sponsored Films*. (San Francisco: National Film Preservation Foundation, 2006), vi.
[4] HD video frame rates currently vary from 24 fps to 60 fps, see Marcus Weise and Diana Weynand, *How Video Works: From Analog to High Definition*. 2nd ed. (Boston: Focal Press, 2007), 121-133.

The arrival in the commercial sector of digital scanning technology, which can mirror both image quality and frame rate of the scanned film, promises to provide a more faithful replication of celluloid film records. Because frame-integrity scanners image each celluloid frame as an image frame and not a video field, the resulting digital file is a much more accurate surrogate of the celluloid object. But as these technologies continue to be driven by the demands of the commercial film market, the tools currently available are skewed to accommodate the segregated workflow of commercial production practices, not the access and preservation needs of archives, libraries, and museums.[5]

For archives, libraries and other cultural heritage institutions facing the need to convert their film collections in order to keep them accessible to researchers, the only option at present is to follow the procedures of the commercial production facilities. This requires the creation of a digital moving image file through one procedure, then the creation of a separate digital audio signal by means of aging analog optical sound readers, and fmally the combination of these two files into one with a careful attention to synching the audio and video together properly.

<u>Proposed Optical Sound Recovery Technology Tool:</u>

Based on successful experiments in 2009 (see "History, Scope and Duration") we propose to develop an open source, royalty free application that can recover the optical sound track from scanned motion picture sound film (see figure 8).

Exemption 4

Figure. 8. An over-scanned image captures content above and below the frame as well as the audio track. Here a variable density soundtrack has been captured with the image.

The proposed software would import existing digital media files (e.g., dpx, avi, mov, etc.) and output a synchronous digital picture and sound file in a variety of formats depending on user preference. It would also allow for the output of a freestanding digital audio file. The software would work with either positive or negative images of color or black & white film stock.

Although MIRC has its own 4K film scanner, the proposed software will work with any scanner meeting three basic criteria, thus maximizing its usefulness. The tool will be designed to process all scans that: 1)

[5] In commercial media production film scanners are used to create digital intermediate (DI) files from which the image content of the finished media is edited. The DI process still typically ends with a celluloid print being generated from the frame-integrity digital file.

contain an optical sound track and image, 2) have stable frame integrity with a minimal overscan, and 3) be of sufficient scan resolution to allow for quality decoding of the optical sound information.

Initial tests have proved the concept, but, broader input from film preservationist and film-sound experts must be sought in order to refine the algorithms so they faithfully reproduce all known single channel optical audio formats and the most common dual-channel formats. To accomplish this, the grant proposes to impanel a board of advisors drawn from film preservation experts across the country and contract with a film-sound professional to assist MIRC and IMI in refining the audio output.

The six-member Board of Advisors (BOA) will be comprised of technical, curatorial, and digital information technology experts and will be charged with providing a critical evaluation of the application's fidelity, practicality and potential. Although the BOA is not yet complete, at the time of submission the following have agreed to serve: Caroline Frick-Page (Curator, Motion Pictures, George Eastman House); and Ralph Sargent (President Film Technologies Inc) and Howard Besser (Director, Moving Image Archiving and Preservation Program , New York University). The BOA will help identify film formats, and digital file formats that should be supported in order for the tool to be most useful to the archival, library and museum communities and they will help establish priorities in the development of the tool. It will also advise as to the collection and packaging of metadata as well as integration with other extant digital preservation workflows.

The BOA will twice gather twice in Columbia, SC to view demonstrations of the software's capabilities and compare its digital sound to analog (see Plan of Work for details).

At such an early stage of development it is difficult to know for certain the capabilities of the initial public release but we expect it to construct digital audio from both variable area (unilateral and bilateral) and variable density formats. It is hoped that most pre-Dolby noise cancellation systems (e.g., push-pull and matted tracks) can be accommodated in the first public release of the software.[6] The Motion Picture and Recorded Sound Division of the Library of Congress has agreed to supplement MIRC's film collection as needed to ensure that all optical sound formats recommended by the BOA are incorporated into the software.

After deployment IMI will provide five years of technical support for the software. Further development of the software beyond IMI's post-grant support period will be the responsibility of the open source development community.

BACKGROUND OF APPLICANT

Moving Image Research Collections (MIRC) at the University of South Carolina (www.sc.edu/library/mirc) is a new unit encompassing the Newsfilm Library, which began with the donation to the University of 11 million feet of Twentieth Century Fox newsreel film. The Newsfilm Library grew to an estimated 20 million feet of archival film and in 2009 with the donation of the Chinese Film Collection from the Chinese Embassy its diverse collections were reorganized to reflect its expanded mission. Total holding are now estimated at 6000 hours, establishing MIRC as one of the largest public film archives in the US.

MIRC's Newsfilm Collections house some of the earliest extant commercial optical sound negatives (all variable density tracks). MIRC's collections also contain a wide variety of optical sound tracks on 35mm

[6] For the progression of early optical sound formats see, Academy of Motion Picture Arts and Sciences, *Motion Picture Sound Engineering*, 3rd ed. (New York: D. Van Nostrand, 1938), 23-43.

and 16mm release prints and camera original television newsfilm. The digital reformatting of archival film has driven much of MIRC's interest in an easy to use and freely available application to decode these valuable sound artifacts.

MIRC's technical staff has expertise in film-to-video transfers and maintains and operates a twenty year-old Bosch BTS FDL 60 telecine. The staff's experience with this quality machine has also reinforced the need to move beyond telecine transfers as the parts and repair expertise for it are rapidly dwindling.

Currently, MIRC is one of the few film archives in the nation to have its own frame-integrity film scanner. The DeMott/Kreines Kinetta film scanner was developed in close cooperation with staff at MIRC to meet the archive's needs. Having both a large and diverse collection of sound film, the experience making broadcast quality transfers of these films via its telecine unit, the capability to create edge-to-edge digital scans in house, and having collaborative research partners from IMI uniquely positions MIRC to carry out the work outlined in this proposal.

The Interdisciplinary Mathematics Institute (IMI) of the College of Arts and Sciences at the University of South Carolina (http://imi.cas.sc.edu/IMI/) serves to foster advanced mathematical research with the potential for meaningful application and to facilitate its transfer to the academic, government and industrial sectors.

The IMI was created in 1992 from a grant of the NSF EPSCoR 1 program and quickly became a leader in developing new scientific techniques for a variety of application areas. The IMI researchers have developed cutting edge algorithms and technology in the areas of data analysis, simulation, and signal and image processing based on sophisticated and innovative mathematics methods.

The IMI has had long and fruitful collaboration with leading research centers both in the US and overseas. The following list contains only some of the lasting and important collaborations: Princeton University (Mathematics, Computational & Applied Mathematics, Electrical Engineering, Chemical Engineering, and Civil and Environmental Engineering), Cal Tech (Computer Science), Stanford University (Computer Science, Computational Harmonic Analysis and Statistics), Rice University (EE Digital Signal Processing), Purdue University (Computer Science, Mathematics), Texas A&M University (Mathematics, Scientific Computing, Computer Science), University of Wisconsin (Computer Science, Anatomy), University of Paris VI (Numerical Analysis), University of Paris VII (Statistics), RWTH Aachen (Applied Mathematics), Naval Air Warfare Center (China Lake), Naval Air Warfare Center (Paxtuxent River), Army Research Lab (Adelphi), Air Force Research Laboratory (Wright Patterson, Eglin).

The IMI has a 20-year history of attracting significant funding from government and industry sources. The IMI has been widely supported by federal and state funding agencies, including NSF, DOE, ARO, AFOSR, ONR, DARPA, NGA, NSA, and NASA. It also has a history of industrial collaborations: Mobil Technology Co., ExxonMobil, Schafer Corp., Radiance Inc., ZeroTree, HydroGeoLogic, Inc., Silicon Graphics, Stardent Computer, Ardent Computer, Intel, NCUBE, and IBM.

In more recent years, IMI faculty have begun to collaborate with other highly-valued centers at the University of South Carolina with research topics ranging from super resolution of scanning transmission electron microscopy imaging (Nanocenter and Microscopy Center) to extremely high speed video endoscopy for laryngeal disorders (School of Public Health).

What IMI brings to the table is a firm theoretical foundation in the area of signal and image processing coupled with a long experience in developing practical algorithms and software.

HISTORY, SCOPE AND DURATION

As the result of a 2006 grant from the Library of Congress to enhance preservation of MIRC's Fox Movietone News Collection, MIRC purchased a 4K frame-integrity film scanner (the Kinetta) from the DeMott/Kreines Films Company. The Kinetta, delivered in 2009, was designed to provide edge-to-edge image scans so that all information content on the film could be preserved in the scanning process. This informational content included the optical sound track when present. Seeing the sound track while the image played led to the obvious question, was it possible to decode the soundtrack using algorithms?

After an initial meeting in July 2009 with Dr. Pencho Petrushev and Dr. Borislav Karaivanov of the IMI, MIRC asked that the IMI provide a proof of concept for a decoder. Over the next five months the IMI provided a series of synched-sound digital film based on scans created by MIRC.[7] These early tests proved that algorithms could be designed to perform the task but that the broad utility of the software would depend heavily on its ability to compensate for a number of variables in the film scans.

The software would presume three basic elements from any scan and then work to compensate for any additional anomalies in the film scan or defects in the celluloid film element. Those basic elements are: 1) frame integrity scan with optical track present within the imaged frame; 2) reasonable vertical and horizontal frame stability and some vertical overscan; and 3) sufficient imaging resolution to provide adequate information data about the sound track.

The scans must be sufficiently wide to capture the full width of the soundtrack area. Partial images of the soundtrack cannot guarantee complete reconstruction of the audio signal as in the case of single sided variable area the highest or lowest amplitudes may get clipped.

The scan sequence must provide a contiguous coverage of the film as sound from missing portions of unknown length and location cannot be recovered. Moreover, a moderate amount of overlap between consecutive scans (up to 25%) is welcomed since it improves the reliability of the algorithms used to match the scans and compensate the shading due to uneven illumination.

The scans must possess resolution that is sufficient for reconstructing an audio signal with the full frequency range of the optical recorder used to create the soundtrack. This requirement is probably the least obvious of the three but is not less fundamental. It can be characterized quantitatively by a rigorous mathematical theory. According to the Shannon sampling theorem[8] in order to completely recover a band limited analog signal, one must sample it at a rate equal or higher than the Nyquist rate that equals twice the highest frequency present in that signal. With modern optical recorders having bandwidth from 20 Hz to 14 kHz, the smallest sampling rate that is guaranteed to capture all stored frequencies is 28 kHz. This translates into at least 28,000 samples per second of soundtrack, or 1167 samples per image frame assuming rate of 24 frames per second. Adding 25% of overscaning, a single scan would need to have height of about 1460 or more lines of pixels.

There is a small but growing body of published research on the use of digital imaging to reconstruct audio from motion picture film. Poetsch et al (2000) "Restoration of optical variable density soundtracks on

---

[7] Versions of these files are available at http://www.math.sc.edu/~karaivan/SoundRestoration/

[8] C. E. Shannon, "Communication in the presence of noise," *Proc. Institute of Radio Engineers*, vol. 37, no.1, pp. 10-21, Jan. 1949. Reprint as classic paper in: Proc. IEEE, Vol. 86, No. 2, (Feb 1998)

motion picture films by digital image processing" formulate a scanning routine specifically designed to facilitate audio reconstruction. [9] Using a one-line, charged coupled device (CCD) camera their scanner captures only the soundtrack area of the film, a single line at a time. Its operation requires very precise manual aligning of the camera with the soundtrack area. On the algorithmic side, they describe a calibrating procedure whose goal is to compensate for uneven sensitivity of the CCD sensors and non-uniform illumination by the scanner's light source. They also address shading due to incorrect positioning of the exposing lamp during the sound recording. Their method was tested experimentally by Hassaïne et al (2009) and found to generate noticeable background noise by making corrective changes that lead to too much variation between the consecutive scan lines.[10] In another work employing the same scanning device, Richter et al (2003) "Localization of faults in multiple double sided variable area code soundtracks on motion picture films using digital image processing" describe a method for localization (but not repair) of multiple variable area soundtracks (also known as Maurer tracks), a rare type of film that may not be incorporated in our initial release.[11] Restoration of variable area soundtracks with significant damage due to scratches, dust, abrasion, or moisture was first addressed by Brun et al (2007) in "Restoration of variable area soundtracks" and later much improved by Hassaïne et al (2009) in "Efficient restoration of variable area soundtracks".[12] Both methods carry very high computational cost, making it difficult to justify their application on films of moderate or better condition where less expensive techniques might be able to accomplish comparable results. Another condition with degenerating effect on the reconstructed sound is the under/overexposure of the film stock. Taquet et al (2008) "Detection and correction of under/overexposed optical soundtracks by coupling image and audio signal processing" model the phenomenon by morphological erosion and dilation, and use simulated film data to validate their technique consisting of two audio-based detection indicators and an iterative correction routine designed to minimize them.[13]

The several papers discussed above originate from two research groups in Europe, one in France and one in Germany. The main acquisition devices used in these experiments are scanners with one-line, charged coupled device (CCD) camera that capture only the soundtrack area of the film, a single line at a time. Similar scanning technologies for motion picture optical soundtracks are being developed by the industry for their own needs. Such scanners are constructed by modifying existing operational telecine equipment while others are proprietary, built for the purpose devices. Both types necessitate a separate scanning pass dedicated exclusively to the sound reproduction.

In contrast, we propose a universal technique that is not confined to any particular scanning device or routine. It can take as input any sequence of digital scans and reconstruct the recorded audio signal. The segregated workflow presumed in the work described above may be acceptable for films undergoing high-end and expensive digital film restoration, but the vast majority of the nation's film record will never be subject to such restoration. With our approach sound reconstruction from the existing digitized materials would be possible without the need for new hardware or reorganization of the present image-acquisition process. One-pass digital preservation of image and sound will reduce the expense of digital

---

[9] Proceedings of the International Conference on Optimization of Electrical and Electronic Equipments, 3 (2000): 793-798.
[10] Hassaïne, A. and E. Decencière, B. Besserer (2009). "Restoration of variable density film soundtracks," *17th European Signal Processing Conference*, 2009: 2589-2593.
[11] Richter, D. and D. Poetsch, A. Kuiper (2003). "Localization of faults in multiple double sided variable area code sound tracks on motion picture films using digital image processing," *Proceedings of the 13th International Czech - Slovak Scientific Conference Radioelektronika*, Brno, Czech Republic, May 2003.
[12] Hassaïne, A. and E. Decencière, B. Besserer, "Efficient restoration of variable area soundtracks," *Image Analysis & Stereology* 28: 113-119.
[13] Taquet, J. and B. Besserer, A. Hassaïne, E. Decencière, "Detection and correction of under/overexposed optical soundtracks by coupling image and audio signal processing," *EURASIP Journal on Advances in Signal Processing*, Vol. 2008: 17 pp.

preservation enabling more films to be preserved in this manner. A review by the University's Intellectual Property Office indicates that this approach is unlikely to infringe on existing patents in this area.

Our preliminary tests found the following issues affecting the sound reconstruction process to be commonly present:

1. Inconsistency of the vertical placemen of the image frame in the form of a slight drift or a sudden jump, making frame alignment more computationally intensive;
2. Uneven lighting, creating inconsistent exposure levels along the optical sound track when viewed over the height of the frame;
3. Disruptive splicing, destroying the integrity and altering the appearance of the optical sound track;
4. Physical damage to or decay of archival film that compromises the consistency of the optical sound track.

An inconsistent amount of overlap from one frame to another requires the software to calculate the exact amount of overlap for each consecutive frame pair (see figure 9). The probability of such vertical inconsistency is significant for archival films for two reasons. Due to age and shrinkage they may not pass through the scanner in an ideal manner. Moreover, camera and printer aperture anomalies present during the exposure of the film may create vertical registration shift, especially after several generations of printing.

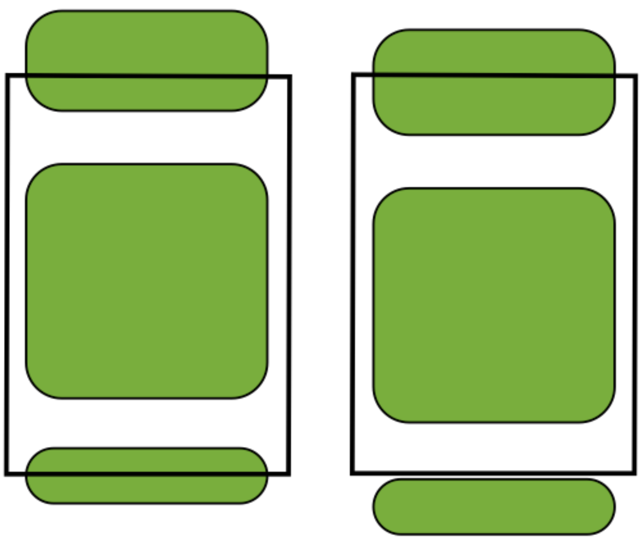


Figure 9. Vertical shift that has been noted to occur in scans after several hundred frames. Here that shift is simulated. The black outlines designate the sensor scanning area. Note that the left scan includes the bottom of the previous frame in addition to the top of the next frame. The right scan shows that a slight shift due to a film or scanning anomaly has moved the position of the central frame down.

An intelligent approach would take into account the fact that the amount of drift is small and use the size of overlap from one pair as an initial guess in the computations for the next pair that could make convergence faster and produce significant savings in resources. In the preliminary analysis we observed occasional fluctuations in the vertical position of the picture frame within the scanned image over a range of 10-15 scan lines. The possibility of jumps in position cannot be ignored and has to be adequately addressed even at a higher computational cost.

The second issue results primarily from inconsistent illumination and light quality throughout the imaged area (see figure 10).

Exemption 4 **Exemption 4**

(a) (b)

Figure. 10: Image (a) shows a single frame cropped from a scan with inconsistent illumination, creating a shading effect in the lower region of the frame. The presence of shading might not be immediately noticeable for an untrained eye but it is revealed indisputably on (b), where the frame from the next scan has its top pmtion replaced by a crop from the bottom of the previous scan corresponding to the same area on the film. This combined frame illustrates that physically the same region of the film appears much darker when it is near the bottom of the scanning window.

Since variable area and density tracks were designed to be decoded through a narrow slit, reading-in effect-one line at a time, variations in illumination did not create unintended sound anomalies. In contrast, frame-integrity imaging of the sound track varies the quality oflight over a much longer section of the audio information and can create anomalies in the imaged sound track. The sound signal in these instances contains loud noise modulating in accordance with the frame rate of the film not present in the original optical track. Coping with the serious challenge posed by inconsistent illumination requires powerful mathematical tools that are discussed below (see, Methodology and Standards).

Disruptive splicing results from the ends of two lengths of film being joined together in a manner not preserving the continuity of the soundtrack image (see figure 11). In this case, the corresponding portion of the reconstructed audio signal is shaped as a gate function causing an audible "pop". The software will discern the difference between the blooping tape (black tape over the track at the splice) and the gap created by the splice itself.

Exemption 4

Figure 11. Disruptive splice with black blooping tape and too large of a gap between film edges. University of South Carolina.

The final issue, physically damaged film, has the most variations. Some damage, such as the vertical scratches running throughout the frame (also known as *rain)* is easily managed. Damage to the optical track from rain has relatively little effect on the sound reconstruction process because the scratches' occurrence over the soundtrack does not alter the shape of the curve only the area under the curve (see figure 12). The change in area can be automatically corrected at the last stage of the sound reconstruction when the level of the extracted audio signal is adjusted relative to the threshold of human perception.

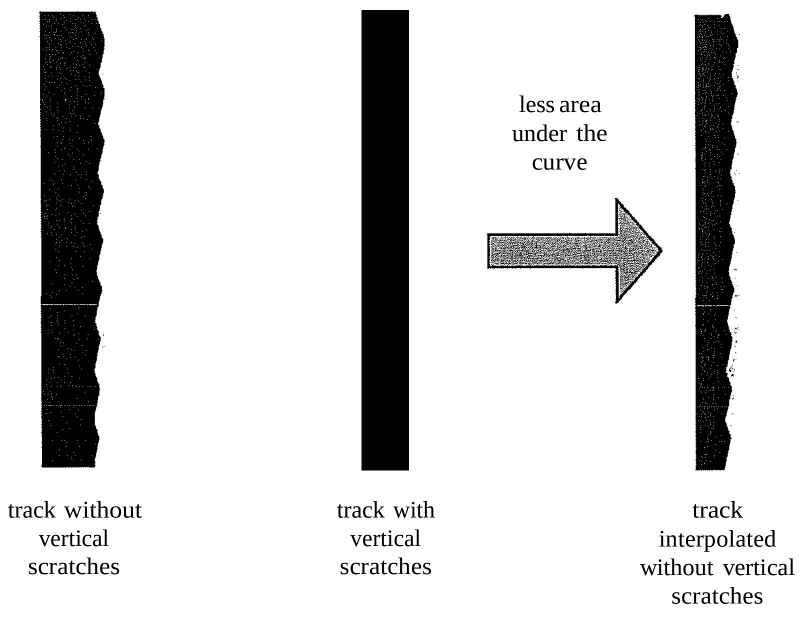


Figure 12: Effect of vertical scratches within a variable area soundtrack. The middle track simulates scratches. The area under the curve is calculated without this data resulting in a smaller area under the curve at right but leaving the shape of the curve intact.

Localized damage to sound tracks from mishandling or decay produce sharp, short peaks or dips in the audio signal, which are heard as dropouts, pops, crackles or a static noise similar to that caused by disruptive splicing. In image restoration, localized scratches are patched over with material borrowed from the neighboring frames under the reasonable assumption that consecutive frames differ very little. Sound is much more dynamic and a similar assumption does not seem to be well founded. For scratches of small area an intuitively better solution is to identify the span of the peak (dip) and replace the audio signal there by a patch taken from nearby and smoothly modified at its ends to match the undamaged signal.

Although image scanning quality might over time reduce the frequency of or eliminate the occurrence of the first two issues, the near term future of affordable digital film scanning is likely to include scanners lacking perfect registration and illumination, especially for public and non-profit archives. Even where ideal scanning conditions are present archival film stock will always generate audio anomalies of the type described by issues three and four.

Our three-year project will result in a freely available software application that can address all four issues described above. After its release, the IMI will provide five additional years of technical support for the software.

METHODOLOGY AND STANDARDS

Existing modes of analog or digital interpretation of optical sound tracks are based on line-by-line analysis of the wave information mirroring the process by which the optical track was created in camera. Sound motion pictures are, in essence, a marriage of intermittent visual information (frames) and continuous audio information (sound track). As long as preservation and digitization approaches view the image and sound information as fundamentally different, film preservation and digitization will remain mired in a two-pass workflow. The film must be processed once to capture the image and then another time to capture the audio signal. Our approach breaks with this model.

We propose to develop a software tool capable of reconstructing an audio signal from a sequence of gray scale scans of overlapping portions along the length of a photographic film. The general process of sound reconstruction is comprised of the following five tasks:

1. Location of the soundtrack band(s) within a scanned image;
2. Conversion of the soundtrack band(s) to a digital audio signal;
3. Determination of frame overlap;
4. Integration of the audio signals from the consecutive frames while correcting for film defects and scanning anomalies.

Accomplishing these tasks requires compensating for the four issues described above (see, History, Scope and Duration) in ways that maximize computational efficiency. Below we will address each of the four basic tasks in more detail explaining the known challenges and limitations.

Locating the sound track within the overlap image is the first challenge. We will develop this capability in two phases. At the early stage the soundtrack band will be located semi-automatically through a guided interaction in which possible boundaries are suggested and visualized by the software with the option to be corrected by the user according to his/her own judgment.

Once this semi-automated phase is complete, we will develop an automated track location feature. In this phase, international standards will form the foundation of a database containing information about film size, location and size of the picture frame, location and width of the soundtrack band, type of recording (variable density, variable area, etc.), type of perforation, frame rate, etc.[14] This database of technical standards will be tested against and refined by scanning samples. Availability of such database will allow us to deploy algorithms for automated soundtrack detection. A scanned image will be compared against each of the templates and their degree of conformity will be measured numerically by a specially designed metric. The numerical ranking created in this fashion will reveal the most likely film format for the tested image, or indicate that none of the formats registered in the database is an acceptable match.

The relative position of the image sensor and the film cannot be absolutely fixed and it changes any time a new film is placed in the same scanner. Hence, the pixel size varies from one scan sequence to another and the standardized parameters such as soundtrack width and location cannot be readily converted from absolute units (inches or millimeters) to pixel indices. Nevertheless, knowing the film format, one can reasonably expect to locate the soundtrack to within several pixels that would be a satisfactory initial approximation that can be perfected by a localized segmentation technique. In our preliminary research we identified several promising classes of algorithms that need to be fine tuned for the task and extensively tested: e. g. (from simpler to more sophisticated) color reduction, flood fill[15], multiresolutional analysis[16], and active contours (also known as "snakes").[17]

In the case of variable-density recording converting the soundtrack pixels to a digital audio signal is a relatively simple task involving horizontal averaging of the gray scale pixel values. For the variable-area formats two options appear equally plausible. By design, a horizontal line from an asymmetric variable-area soundtrack is meant to consists of two intervals, one purely black and the other purely white, with their lengths adding up to the soundtrack width.

One possibility is for every fixed pixel row to determine the positions at which the black area starts and ends by looking for the maximum contrast between adjacent pixels on that row and accounting for noise in the process. Disregarding their actual colors the pixels between the two found positions are assumed to be black and the rest of the pixels white. The value of the audio signal for that row is calculated as the ratio of the length of the black interval and the soundtrack width, both measured in pixels. For a typical film 35mm film scan the soundtrack width is in the vicinity of 200 pixels implying that such audio signal will be coarsely quantized since each entry is limited to one of approximately 200 values. This disadvantage is partially offset by the immunity of the approach to variations in illumination, which is practically eliminated when pixel values are altered to black or white.

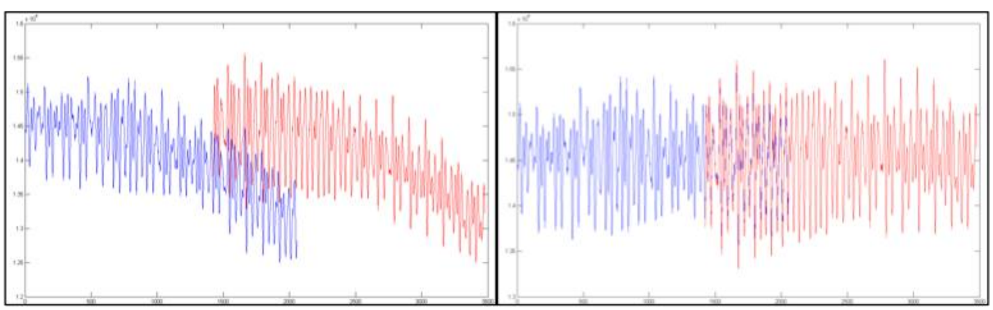
The second option is to apply the same routine as in the case of variable-density soundtrack. On each pixel row the true boundary between the dark and light areas of the soundtrack will occur inside a pixel. That pixel will have a gray value obtained by averaging the colors of the fine film grains inside of its square area of which some are dark and others are light. The advantage of this method is in producing a

[14] International Standards Organizations, International Classification Standard 37.060.20 provides most of the relevant ISO standards for motion picture audio sound tracks. The Society of Motion Picture and Television Engineers also publishes standards relevant to this work, including SMPTE 40-2002, SMPTE 48-1995, SMPTE 111-2001 and many others.
[15] Soille, P. (1999). *Morphological image analysis: principles and applications*, Springer-Verlag.
[16] Mallat, S. (2008). *A wavelet tour of signal processing, 3rd ed.*, San Diego: Academic Press.
[17] Kass, M. and A. Witkin, D. Terzopoulos (1987). "Snakes: Active contour models," *International. Journal of Computer Vision* 1: 321–331.
Singh, A. and D. Goldgof, D. Terzopoulos (1998). *Deformable models in medical image analysis*, Alamitos, CA.: IEEE Press.

sound signal that is much less quantized. Its shortcoming is in its sensitivity to shading as the unmodified pixel values directly enter the calculations.

Since over-scanned images are required to capture all of the audio information, the resulting vertical overlap from one frame to another must be determined and integrated into computations. Image registration aims to solve the general problem of matching two images that partially contain the same unknown scene viewed from different, unknown perspectives. Ignoring the high computational cost we could have chosen to utilize an established method from the vast amount of literature available in the area with many sophisticated algorithms well-described and extensively tested.[18] Instead, we would opt to develop a highly specialized algorithm that takes advantage of the rich specificity of our matching task and allows us to realize gains in both efficiency and accuracy. Reasonable guess for the amount of overlap between two consecutive scans can be derived from the overlaps found in the previous pairs of scans except when a sudden jump is present. We contemplate that refinement of this guess into a more accurate estimate can be achieved using the fact that the Fourier transform of a shifted signal is a complex multiple of the original signal's Fourier transform, and the phase of the complex multiplier reveals the amount of the shift. Two smaller images of the guessed size are cropped, one form the bottom of the first scan and another from the top of the next scan. The Fourier transforms of those images are computed quickly and inexpensively using the Fast Fourier Transform (FFT) algorithm. The shift estimated from the pointwise ratios of the Fourier transforms is used to compute the amount of overlap.

Finally, forming a coherent audio signal out of the pieces extracted from each scan is considerably obstructed when the imaged frame and track are generated with inconsistent illumination. The quality of sound will be much compromised unless the effect of this inconsistency is properly factored out. Figure 13, (a), illustrates the audio signals from two consecutive scans heavily affected by shading. In this case, inconsistent illumination of the imaging area created an increasingly darker region in the bottom half of each frame. This gradual darkening of the image is represented in the sinking of the graphs in 13 (a). Had the illumination throughout the image been uniform, the two graphs would have coincided on the common part of their domains.



(a) (b)

Figure 13: (a) Audio signals from two consecutive, not calibrated scans possessing heavy shading toward the bottom. Ideally, on the middle portion of the window the two signals would coincide as they both represent the same portion of the soundtrack on the film. Due to the shading the second half of the blue signal is gradually lowered. Combining the blue and red signals as they are produces an aggregate signal with a sharp jump at the junction which is clearly heard as a short and pronounced cracking noise. (b) The same audio signals properly compensated for the shading match significantly better on the interval where they overlap. The aggregate signal does not have sudden jumps and the junction is audibly unnoticeable.

[18] Brown, L. (1992). "A survey of image registration techniques," *ACM Comput. Surv*. 24:325-376.

Generally, there are two approaches to attack that problem. One is to remove the shading from the whole scans and use the cleaned images to reconstruct an audio signal. The other approach is to reconstruct a distorted piece of audio from each raw scan, correct each piece for the effect of shading, and blend them together in one uninterrupted audio signal. The second approach appears more appealing to us as a more efficient solution. We envision the per-piece correction for shading as an optimization problem. Assuming that due to the shading the local averages of both signals decay as the same polynomial of a fixed degree, we would find the polynomial's coefficients that minimize the root mean square (RMS) error between the two signals with properly compensated decay according to the sought polynomial. For the two signals from figure 13 (a), the conceived correction is illustrated on figure 13 (b).

Identifying sound defects caused by disruptive splicing or physical damage to the film material can be sought in both image and sound domain with either approach having its advantages and shortcomings. To employ the strengths of both approaches we propose a hybrid technique that benefits from the cost efficiency of the former and reliability of the latter. Being one dimensional, an audio signal can be inexpensively examined for patterns that rarely occur naturally in sound. For instance, gate functions formed by disruptive splicing are revealed by the local extrema of the signal's convolution with an appropriately designed, finitely supported mask. To eliminate the false positive indications, the suspected segments from the signals can then be further analyzed on the image side where more information is available. For defects of relatively short duration the damaged audio signal can be replaced by a patch taken from nearby and smoothly modified at its ends to match the undamaged signal. In case of larger faults sound restoration may not be feasible but their effect can be mitigated by properly silencing the noise over the correctly identified span of the defect.

Because the primary function of this software is to preserve the motion picture audio and image in a digital form satisfying the demands of scholars, curators, and artists, critical evaluation and quality control will be ongoing throughout the grant period. At IMI, researchers will evaluate the mathematical representation of the sound waves generated by the software. At MIRC, the digital sound and image files created by the application will be evaluated against the Betacam SP master videotapes produced by the staff on the Bosch BTS FDL 60 telecine. These two types of testing will be ongoing through the project. After sufficient development in year one, the software and samples will be sent to a professional audio engineer for critical evaluation. Meetings of the Board of Advisors at National Audio-Visual Conservation Center will provide another degree of evaluation.

PLAN OF WORK

Year one of the grant will have as its principle goal devising effective solutions to the four issues explained above (see, History, Scope and Duration): 1) vertical frame shift, 2) non-uniform illumination within the scan, 3) disruptive splicing, and 4) physical damage on the film. By year's end a stable, working prototype of the application will be operational. Initial test will be sent to Bob Heiber and Chace Audio for preliminary assessment.

<u>Year One Activities</u>:

- Contract with professional motion picture audio consultant, Robert Heiber of Chace Audio by Deluxe (Cooper)
- Impanel Board of Advisors and seek initial input from its members (Cooper and Wilsbacher)
- In consultation with audio engineer and relying on SMPTE standards MIRC will develop a database of optical sound formats for use in the software (Wilsbacher)
- Hire temporary staff to provide Kinetta scans in accordance with the grant's needs (Cooper and Wilsbacher)
- Purchase computational equipment for IMI researcher (Cooper)
- Purchase equipment required by MIRC (PC, monitor, headphones, etc..) (Cooper)

- Prepare and submit materials to Dr. Karaivanov and Robert Heiber (Wilsbacher and technician)
- Identify and implement computational solutions to address the issues identified by preliminary testing (Karaivanov)
- Evaluate computational solutions on the basis of mathematical efficiency and audio fidelity (Karaivanov and Petrushev)

Year two will start with a two-month period of intensive testing with scanned films of various formats and degree of preservation in order to identify unaddressed issues and possible weaknesses in the year- one implementation. Bob Heiber and Chace Audio will assess submitted examples against industry standards. Heiber's recommendations will be assimilated into the core computational solutions. Year two will end with a meeting of the Board of Advisors and key grant personnel in Columbia, SC. The convened Board will see a full demonstration of the digital workflow, evaluate the results of the software and make recommendations about the design of the Graphic User Interface for the tool.

<u>Year Two Activities:</u>

- Prepare and submit materials to Dr. Karaivanov and audio consultant (Wilsbacher and technician)
- Develop automated track location feature (Karaivanov)
- Finalize computational solutions for all identified issues based on feedback provided by audio engineer (Karaivanov and Petrushev)
- Schedule and convene BOA meeting at Columbia, SC (Cooper)
- Summarize the findings of the Board of Advisors meeting and note any further enhancement of the tool required in year three (Karaivanov and Wilsbacher)

Year three will begin the migration of mature core code into an alpha version of a GUI and finalize all functionality to be included in the first public release of the software. Dr. Karaivanov will revise algorithms and computational solutions based on the results of the first BOA meeting. The BOA will evaluate the alpha version of the software and provide final critical feedback on the fidelity of the digital audio construction. The BOA's divers experience will ensure that the initial release of the software meets the needs of a broad array of projected users. Midway through the year a beta version of the software with GUI will be made available to a group of beta testers. Feedback from the beta tests will be incorporated into the first public release of the software, which will coincide with the end of the grant. Drs. Karaivanov and Wilsbacher write the required white paper in the final month of the grant.

<u>Year Three Activities:</u>

- Hire part time programmer (Petrushev and Cooper)
- Prepare and submit materials to Dr. Karaivanov audio consultant (Wilsbacher and technician)
- Develop and test in-house alpha version (Karaivanov)
- Optimize performance of the code through reduction in the usage of computing time and memory resources without compromising any of the existing functionality (Karaivanov)
- Schedule and Convene Board of Advisor's meeting at Columbia, SC (Cooper)
- Develop beta version (Karaivanov)
- Write manual for beta version (Karaivanov, Petrushev and Wilsbacher)
- Attend scholarly conference(s) (Karaivanov)
- Identify beta testers (Cooper and Wilsbacher)
- Distribute beta version and collect feedback (Karaivanov and Petrushev)
- Develop first public release (Cooper, Karaivanov, Petrushev and Wilsbacher)
- Advertise release via list-serv and other professional communication venues
- Write white paper as required by grant (Wilsbacher and Karaivanov)

STAFF

Dr. Greg Wilsbacher, Curator of Newsfilm Collections, (PI) has worked extensively with DeMott/Kreines Films in the development of MIRC's scanner. Not only is he able to operate and maintain the scanner he also instructs undergraduate interns to operate the scanner. He has been instrumental in developing MIRC's digital media capabilities. In 2009 he was invited by the Library of Congress's Office of Strategic Planning to participate in its "Preserving Digital News" strategic planning symposium.

Dr. Mark Cooper (Co-PI) is Associate Professor of Film and Media Studies and English and Interim-Director, Moving Images Research Collections. Dr. Cooper is the author of two scholarly books, *Love Rules: Silent Hollywood and the Rise of the Managerial Class* (2003) and *Universal Women: Filmmaking and Institutional Change in Early Hollywood* (2010).

Dr. Pencho Petrushev (Co-PI) is Director of the Interdisciplinary Mathematics Institute and a professor in the Department of Mathematics, USC. His areas of expertise are Harmonic Analysis, Approximation Theory and their applications, in particular to signal and image processing. Petrushev has worked on various project funded by NSF, DARPA, NIH, NGA, ARO, ONR. Most recently he has been working on a project with Dimitar Deliyski (Director of Voice and Speech Laboratory, Communication Sciences and Disorders, Arnold School of Public Health, USC) in the area of laryngeal high-speed videoendoscopy (NIH/NIDCD: grant R01DC007640 "Efficacy of Laryngeal High-Speed Videoendoscopy", 2007 - 2012). Another recent project developed by Petrushev was in the area of multiscale data representation on the sphere and on the ball based on newly created wavelet type systems, called "needlets", with application to geopotential modeling (NSF-DMS 0709046 "Highly effective representations for surface and solid spherical studies", 2007-2010). [http://www.math.sc.edu/~pencho/]

Dr. Borislav Karaivanov, Research Associate Professor (Co-PI), Interdisciplinary Mathematics Institute, has conducted the preliminary research providing proof of concept for this project and formulated its main technical components. For the last 10 years he has been engaged in innovative research developing efficient algorithms and software for variety of practical signal and image processing problems. His work was supported by federally funded grants including projects solicited by ARO, NGA, DARPA, and ONR.

Robert J. Heiber, Vice-President Audio/DDM, Chace Audio by Deluxe, will provide professional audio consulting services.

Temporary Scanning Technician (to be hired).

Part time programmer, 3rd year only (to be hired).

DISSEMINATION

Our proposal seeks to develop open-source software to serve a truly global community of film archives. To ensure we are addressing all the core issues of this community presentation of this work at conferences is essential. Drs. Wilsbacher and Karaivanov will jointly submit a paper to the triennial gathering of audio-visual conservation specialists, the Joint Technical Symposium (JTS) to be held in 2013. Organized by UNESCO and the Coordinating Council of Audiovisual Archives Associations (CCAAA), JTS regularly draws on participants form nine film and audio preservation organizations. The location for this conference has not been published but it is often held outside the United States. While the JTS will be a relatively expensive conference, no other gathering of technical film and sound preservationists occurs anywhere in the world. The opportunity to discuss the research and demonstrate the software with an international gathering of experts should not be missed.

A web site created and maintained at the facilities of the Interdisciplinary Mathematics Institute (IMI) at the University of South Carolina will provide general overview of the goals and purposes of the project as well as updates on its progress. It will feature both information educating the general public on the scope and capabilities of the product being developed and more technical notes aiming at the interested specialists in the field of sound motion picture digitization and restoration.

Upon completion of the project the release version of the software will be freely available upon request to all parties with non-commercial interest in the software. Potential users will be required to identify themselves and their affiliations, and specify the intended applications for our product and the operating system(s) under which the program will be used. Collected information will serve only for producing usage stats and disseminating information about updates and new releases.

The software's functionality will be maintained by IMI for the major operating system platforms as they emerge in accordance with the interest in it demonstrated by the targeted communities. IMI’s support will last for a period of 5 after the termination of the grant.