Proceedings Paper:
Schofield, Guy Peter, Green, David, Smith, Thomas et al. (2014) Cinejack: using live music to control narrative visuals. In: Proceedings of the 2014 conference on Designing interactive systems. ACM, pp. 209-218.

Cinejack: Using Live Music to Control Narrative Visuals

Guy Schofield, David Green, Thomas Smith, Peter Wright, Patrick Olivier
Culture Lab, Newcastle University, Newcastle Upon Tyne, United Kingdom
{g.p.schofield, d.p.green, thomas.smith, p.c.wright, p.l.olivier}@ncl.ac.uk

ABSTRACT
We present *Cinejack*, a system for directing narrative video through live musical performance. *Cinejack* interprets high-level musical content from live instruments and translates it into cinematographic actions such as edits, framings and simulated camera movements. We describe *Cinejack*'s technical development in terms of a novel and highly pragmatic approach to interface design, where the affordances of users' own musical instruments are used as controllers through an interpretive interaction scheme.

Author Keywords
Music technology; practice-based research; research in the wild; interdisciplinary design; digital arts

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous. H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)

INTRODUCTION
Cinejack is a multimedia performance system which enables musicians to exercise flexible and expressive real-time control over narrative movies through live music. *Cinejack's* various modules translate musical meaning and expression from users' instruments into cinematographic actions such as edits, simulated camera movements and changes in composition. The resulting visuals can either be used to augment live performances (as is increasingly common in rock and electronic music shows) [1] or to support a new type of performance medium: a form of cinema where music controls, rather than supports film.

In the production of Hollywood cinema, the soundtrack is often accorded a supporting role in relation to the image track [2]. *Cinejack* affords a reversal of this relationship, enabling the user to explore paradigms where the expressive and narrative affordances of music and live performance directly drive supporting visual content. Conceived and developed by a multi-disciplinary team comprising programmers, musicians and film-makers, *Cinejack* is intended to link narrative visual content and live musical performance in real time without the need for a dedicated VJ or programmer by connecting existing conventions and standard approaches from cinematography and music composition.

Cinejack's design is situated within a larger project, the goal of which is a comprehensive set of tools and techniques for musicians and visual artists. Central to the research is an in-depth and long-term exploring how musicians engage with their instruments, their audiences and each other 'in the wild', through the co-creation of a set of performances and tools. Taking as its starting point, the first authors' own practice in an electronic rock band, it has relied throughout its development on generating design insights through different collaborative creative processes.

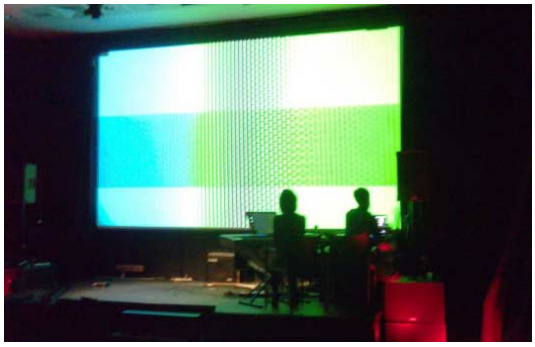

Figure 1. Cinejack in use in performance

Cinejack's design took shape through 3 distinct but inter-related phases. The first involved a scoping project realized entirely by the authors, which was intended to establish both initial conditions for the design of a tool and to frame later discussions with musicians. The second involved the application of lessons learned from this project to the production of bespoke systems for another musician in order to establish critical distance while still maintaining a position within the design space. The third phase involved collaboration on equal terms, as the designers put together an audio/visual production with another musician. The reason for these multiple links to practice is simple: by establishing perspectives from different situations within the design space, we hoped to facilitate creation of a tool which takes into account the context in which musicians practice their art, from rehearsal room to stage and which can therefore be quickly and seamlessly integrated into musicians' performance practices.

The research explores a highly pragmatic form of interface design where familiar interfaces are repurposed for new practical and expressive ends. Rather than requiring of the user that he or she begins to learn to operate a new interface, *Cinejack* relies on the established expressive

DIS '14, June 21 – 25, 2014, Vancouver, BC, Canada

http://dx.doi.org/10.1145/2598510.2598520

affordances of the user's own musical instruments, capitalizing on his or her own skills and developing new controls through already highly developed sets of conceptual and haptic vocabularies.

RELATED WORK

The development of *Cinejack was* driven in part by a lack of available tools that allow musicians to control visuals with minimal, conscious input and with meaningful and aesthetically interesting results. A range of dedicated software packages are available for composing live visuals, such as *Isadora, Motion Dive and Resolume,* however most deal only with fairly low-level musical features such as pitch and amplitude, offering simplistic functions such as beat-matching and triggered responses to hardware MIDI controls. Intended mainly for VJs (artists who compose video in real time in nightclubs and concerts), these tools typically allow rapid composition of video clips against an audio track and most come with banks of filters with which to alter and enhance them, in the manner of a real-time video-editing suite.

For musicians wishing to use video projections in their live shows, many of these tools are inappropriate, due to their mode of operation. Most are intended to be 'played' manually during the performance by a dedicated operator, which in the case of most musical ensembles assumes the addition of another performer to the work. Although this may be perfectly appropriate for VJ practice, the addition of another person to the creative and structural makeup of a musical ensemble is hardly trivial especially in the socio-economic context of a professional music practice where logistical (not to mention economic) constraints are often severe.

Currently, musicians and performers using complex visuals, often rely on building their own systems using visual programming environments such as *Max/MSP/Jitter*, *VVVV* or *PD*. Although these platforms offer the potential for developing complex and visually interesting interactions between music and visuals, this approach requires substantial scratch-building and a high level of technical ability on the part of the user. None of the aforementioned packages contain ready-made setups that are able to glean any expressive or symbolic meaning from the music they analyze, consequently they cannot access the powerful expressive and symbolic connections that music and the moving image can invoke.

Novel Interactions through Familiar Interfaces

In designing a tool to orchestrate audio/visual shows, we looked carefully at work from the growing research community developing New Interfaces for Musical Expression. We considered approaches using novel controllers, such as Bongers and Harris's Video Organ [3] project, which featured an interaction scheme using modular physical controllers which can be patched together to create flexible audio/visual tools. We realized that a similarly modular approach in the design of *Cinejack*'s software could allow the tool to be flexible whilst also producing complex and interesting results.

We also considered augmenting existing instruments, following the example of projects like the IRCAM Augmented Violin [4] or Overholt's Overtone Fiddle [5]. In these cases, instruments' sound generating properties are extended via the addition of extra electronic sensors: accelerometers measuring the movement of the bow and the instruments' body, electronic pickups etc. The authors suggest a range of ways in which these extensions could be used for new musical compositions and playing, which could be reasonably extrapolated to the control of video.

In the context of professional music practice, applying approaches like these to the development of a tool for skilled musicians raises a number of challenges. Foremost among these is the amount of cognitive 'bandwidth' available to musicians engaged in the act of performing. Newton and Marshall's work [6] with augmented instruments surveyed musicians working with modified versions of familiar instruments and found that even by adding new features to these familiar interaction surfaces, musicians could quickly become overloaded, with detrimental effects to their practice.

An additional problem related to virtuosity. As musicians working in a multi-disciplinary research environment we had experimented previously with a number of tangible musical interfaces from REACTable [7] to experimental gestural interfaces and were acutely aware of the investment in time required to gain familiarity and competence with even a simple musical interface. In the context of designing for real-world musicians, this seemed unnecessary, as professional performers already have, by definition, well-developed skills in exploiting the affordances of their own instruments. In considering musicians' relationships with their instruments, we referred extensively to Magnusson and Mendieta's 2006 survey [8] in which over two hundred musicians from cellists to live-coders were interviewed concerning the expressive and practical affordances of their chosen instruments, noting especially the importance placed by many musicians on familiarity and 'depth'.

Having considered these approaches, we realized that the affordances of users' own instruments might be extended by using them to not only generate sound but also to simultaneously control video. Instead of forcing musicians to learn to operate a new and unfamiliar set of controls, the audio output of their own instruments could form the main interface to the system, not through direct 1-1 controls but through an interpretive interaction scheme which could reason over musical meaning. In this way, the full expressive properties of the users' skills in playing these instruments could be fully exploited.

Music and Visuals: Connecting Vocabularies

As well as the theoretical underpinnings of its design,

Cinejack relies on common ground between devices and conventions from a number of different domains. Both contemporary visual media and music rely on a wealth of commonly understood techniques and devices to generate not only strong emotional and visceral responses but cultural and symbolic associations. The semiotic structures used by artists working in film, art and music involve complex cultural associations, whereby cinematic or musical actions, events and subjects often refer to wider narrative ideas [9]. *Cinejack*'s design explores the idea that many of these conventions could function independently of the medium from which they originated and might therefore be transferred between disciplines. A key part of *Cinejack*'s design process involved exploring whether musicians would be able to exploit this phenomenon in live scenarios and how effectively this could be supported by our tool.

In considering the conceptual and technical details of the interface between performance and video, we explored theory from cinema and music composition. Modern film-makers in the Hollywood tradition work largely from a palette of conventions, developed over more than a century of experimentation. Detailed character relationships, dramatic tension and atmosphere are all routinely imparted to the viewer through cinematographic techniques as much as the performances of the actors. In considering possible rule sets for *Cinejack*'s output, we used standard texts by Arijon [10] and Katz [11] to establish ways to relate cinematographic actions to narrative dimensions. Our selection of these rules was largely based on classical Hollywood approaches, chosen solely for their widespread use and consequent legibility by audiences. These decisions did not represent a particular ideological or artistic commitment on our part but merely allowed a set of starting points which other practitioners could either use or subvert.

Although the music of films has not been subjected to anything approaching the level of theoretical attention directed toward cinematography [12], a number of composers and theorists have helped to establish and identify a number of techniques and devices which are commonly used to support or subvert meaning in the film. Modern audiences are highly educated (often without being aware of it) in the languages of film music: being able to read the narrative structure of the score almost as clearly as the plot of the film [2]. We began the design of *Cinejack* using a selection of these conventions: again, as a starting point informed initially by the requirements of our own creative work.

Interactive Cinema

Practitioners from a number of disciplines have experimented with real-time control of cinematic presentations since the late twentieth century. Lev Manovich, in his *Soft Cinema* project, experimented with databases of pre-rerecorded video clips to create multiple narratives from the same dataset [13]. Research projects such Canadian Film Centre's Interactive Narrative Feature Program have led to the first interactive film productions, relying on recombining pre-shot material [14]. Our implementation of *Cinejack* relies heavily on these approaches; recombining video files in real-time according to live inputs.

Research in Interactive Storytelling yielded a number of useful approaches to interactive cinematography and the structuring of mediated narrative presentations. Work by members of the IRIS network, involves applying cinematic conventions to real-time camera control in 3D environments, allowing a cinematic solver to take into account aesthetic and compositional factors as well as dramatic and diagetic concerns when shooting a dramatic sequence [15]. This approach provided a useful framing of features which a system could reasonably deal with and use expressively, avoiding overly open-ended schemes.

Design for Performance Domains

In considering how to translate musical meaning and expression in performance we referred extensively to the work of Thompson et al. around displays of affect in rock and blues performance [16], considering carefully which types of musical gesture to include within the scope of the system. However, as practitioners, we realized that to concentrate solely on the musical outcomes of a performance was to ignore the context of those outcomes: the details of rehearsal and preparation, the technical limitations and affordances of instruments and technology and the social makeup of a performing group. The impact of any of these might affect the usefulness, or creative potential of a tool for visuals.

In designing tools for artistic performance of any kind, it was clear that we would need to work closely with our target users to fully understand the domain. By using co-creation and collaboration with artists as design tools, this project draws substantially from work by Hook et al [17]. Hook's work with VJs used video of participant artists at work as a focus for discussion and self-reflection, generating insights into the co-creation of a performance system. In the development of *Cinejack* however, we added a novel element in which we, as practitioner/designers, placed ourselves fully within the design frame, experiencing the benefits and risks to our own practice of each element of the design in alongside our participants as we gradually pass design iterations back and forth from the test-bed of our own practice to those of others. This approach differs from conventional design relationships as our discussions with each participant are between fellow practitioners, with common and explicit motivations and a common stake in the outcomes of the process.

Finally, in planning how to evaluate the *Cinejack* project, it was necessary to consider very carefully the interactions taking place both on and off the stage. We used approaches by Reeves, Benford et al [18] and Sheridan et al [19] to categorize the parties involved and think clearly about their roles, in terms of their relationships with each other and

their audiences and their position within a performance frame. As we were considering both ensembles and solo musicians as potential users of the tools we were developing, it was necessary to think not just about connections between performer and visuals, but performer and performer, performer and instrument and performers and audience.

DESIGNING *CINEJACK*

From the outset of the research, we were very aware that laboratory-based approaches to designing performance tools might well overlook crucial aspects of the performance experience as it is lived. Buxton jokes that there are three specifications for design of interfaces: standard spec, military spec and musician spec, of which musician spec must be the most robust and reliable [20], highlighting the demands for stability, robustness and fitness for purpose placed on technology for stage environments. For this reason we were determined to situate the design process 'in the wild' as much as possible. In designing for other musicians, we felt that it was integral to the success of the work to not only take into account the unpredictable, high pressure, sensorially confusing environment of the concert stage but also to consider the ways in which musicians prepare their shows, through jamming, solo practice and group rehearsal.

As discussed, the problem of limited cognitive and haptic 'bandwidth' means that while performing we are usually fully engaged in the act of performance and have little attention left to give to control of visuals and lighting. As well as using musical gestures as the mode of interaction it was also crucial that any other technological components of the project were kept as streamlined and 'hands-off' as possible to avoid unnecessary distractions or breaks in the flow of the performance. At the outset of the project, our own equipment was comprised entirely of off-the-shelf instruments and software. We determined that if other musicians were eventually to use *Cinejack,* we must take care to ensure that the software not only met their needs or aspirations in terms of augmenting performances but also fitted easily into the technical infrastructure of their practices.

The demands of our own performance practice provided us with certain constraints from the outset in the development of *Cinejack,* which we thought reasonable starting points to discuss while working with other musicians. We identified four requirements for the tool with which to usefully frame the design.

- It must enable musicians to control narrative visuals while also engaged in the act of performing.
- It must be able not only to integrate easily with practitioners' existing toolset but to fit into their creative practice as a whole with as little disruption as possible.
- It must allow for differences in style and content, accepting a wide range of mappings and configurations depending on the user.
- It must be able to add significant value to their performances, either in terms of aesthetic satisfaction or in more concrete ways, such as extending the range of performance venues open to them.

The first of our design constraints was relatively simple to fulfill given the recent proliferation of programming environments available to musicians working with lighting and visuals. Max/MSP/Jitter was selected as a platform for developing the project for several reasons. Firstly, it offered a flexible programming environment well suited to audio/visual projects, allowing real-time prototyping and debugging. Max's graphical interface also allowed rapid communication between members of the design team and participant musicians.

Over the course of developing the work, we used 3 distinct but inter-related approaches, explored through 3 separate performance projects. The first of these, *OverWatch*, was written and developed by the authors as a framing exercise to explore a wide range of possible interactions and was tested within the context of our own practice. *OverWatch* was a narrative audio/visual piece involving control of a real-time 3D engine through a live musical score. Presented as a Silent Cinema piece, the authors performed a live electronic score beneath a projection screen with the output of their instruments triggering animations, adjusting lighting and editing between multiple virtual cameras. The narrative was a simple science-fiction tale about a deserted world overwhelmed by CCTV. This played out over 4 scenes, controlled by 4 different musical movements, each of which featured a slightly different interaction scheme. In each scheme, the system responded to different controls, mainly in the form of simple motifs, recognized as combinations of notes.

The resulting performance was successful against the terms of our practice: it was novel, visually striking and – equally importantly – satisfying to perform. However there were a number of striking limitations to the underlying software which suggested alternative approaches. We found that the sheer number of possible outputs: animation, cameras, and lighting: led to a somewhat arbitrary and confusing interaction scheme. Audience members interviewed after the concert were largely unable to determine exactly how the music had affected the film. We determined that a more tightly framed system, where editing and camera controls were the focus, would be more manageable and legible and therefore capable of more expressive results.

OverWatch was designed and implemented specifically for our own use and took the form of a series of tightly connected Max modules. This meant that the system's inputs and outputs and the mappings between them, although highly appropriate for our own practice, were somewhat rigid and specific, restricting our playing and

largely ruling out improvision. We realized that flexibility and rapid reconfiguration would need to be built into the design if other users were to integrate it into their practices: consequently, we structured the continuing research to include very different use cases.

The second and third phases of the research involved 2 overlapping projects, intended to explore 2 of these possible scenarios. The second project, *Golden Shroud* was developed in conjunction with another musician, developing a simple, streamlined system to produce visuals to support their own concerts. The third project, *Orientation*, was a collaboration on equal terms to develop a fully narrative 'Silent Cinema' performance, requiring a far more complex interaction scheme. By designing a single tool to satisfy both these sets of requirements, we ensured the flexibility of the final design.

TECHNICAL DESCRIPTION

Hollywood convention offers a range of potential mappings of musical event to cinematic action: for example, linking minor modes to slow zooms, in order to emphasize dramatic tension. Through working with several musicians during the development of *Cinejack*, we realized that a standardized mapping scheme was not appropriate if the final tool was to be used in a range of creative practices. We therefore designed *Cinejack* as a set of modules which could be linked together according to the user's specific project. Each of these modules, built in Max/MSP/Jitter, fulfills a particular function, such as cutting between video tracks, or allowing monitoring of pitch or tempo. Changes in amplitude, particular phrases and motifs and accidentals outside the key of the music can all be applied, via these monitoring modules to modules that perform cinematic actions such as pans, zooms, reframes, cuts and dissolves. In each case, mappings of musical event to cinematic action can be easily adjusted, substituted or inverted. In addition, each module is designed to be connected as required with the minimum of effort.

The design of the modules attempts to maintain consistency of function, documentation and visual style. Max has a presentation mode, whereby with a single click, patchers can be reduced to just the visual elements required for performance, hiding unnecessary objects and patchcords. *Cinejack*'s custom modules are already prepared for this and on launch of presentation mode, reduce themselves to the bare minimum of information needed for a performance, hiding any distracting elements. *Cinejack* is designed for users who are expert in music but not necessarily trained in film-making, or vice versa. To support these users, each module is supplied with documentation describing its function but also its possible uses including examples from cinema and suggestions as to how the might be replicated.

Description of Modules

Microphones, MIDI controllers or live instruments are connected to a computer running *Cinejack* via any audio or MIDI interface. MIDI, rather than OSC was chosen as the main protocol for communication between *Cinejack* and other devices, solely for its ubiquity and compatibility with hardware. Resulting MIDI or audio inputs to *Cinejack* pass first through 1 of 2 modules. *AudioListener* takes audio signals from the soundcard and passes them first through stereo gain controls (which can be quickly mapped to a MIDI controller for easy adjustment) and then to a pitch-tracking object which extracts fundamental pitch, amplitude in dB and attacks (measured as sudden peaks in amplitude). Its pitch and amplitude outlets are configured to send information to the inlets of any other modules listening. Audio can also be routed through Cinejack via the commonly used Rewire protocol, allowing control of audio inputs and outputs from another music application. The *MIDIListener* module, functions in a similar way to *AudioListener*, but uses a configurable connection to a MIDI port and a MIDI velocity.

Cinejack relies on a flow of information from musical input to video output (figure 2.). All *Cinejack* setups therefore comprise at least 2 modules, either an *AudioListener* or *MIDIListener* module and at least one *Cut/Fade* which controls the separate *Cinejackplayer* video module and connects to a video projector. Modules connected in between, interpret particular musical structures, or make specific video transform actions.

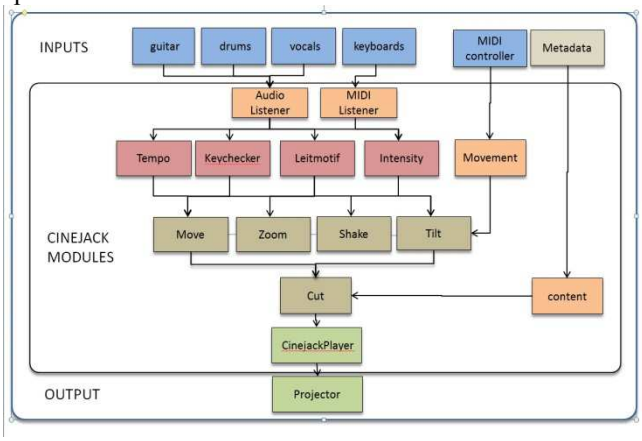

Figure 2. Some of *Cinejack*'s modules and their possible connections.

Many of *Cinejack*'s monitoring modules rely on pitch information provided by MIDIListener or AudioListener modules. Since the Common Practice Period (1600-1900), Western music has been dominated by the tonal system [12] a hierarchical approach to pitch, that places an emphasis on each note's relationship to a tonic or central pitch. In the vast majority of western popular and classical music, major and minor scales are used to determine which notes are included within the structure of the piece (i.e. in the same key) and which should be considered outside it. Consonance or dissonance within this system is a powerful tool for suggesting harmony and discord, tension and release in a narrative. Relationships between pitches can be

used to create specific effects or to refer directly to spatial elements of a drama, for example, slowly rising glissandos can generate tension while descending notes.

In *Cinejack*, these effects can be exploited through a number of modules. The most simple of these: *Keycheck* is a simple look up table to identify notes outside the key signature of the music. Accidentals are reported as simple bang messages and can be used as measures of harmonic discord and, through output modules, mapped to effects to suggest instability or tension such as tilts or small camera movements.

More complex pitch-based modules include *Leitmotif*. Leitmotifs are distinct musical phrases, recurring throughout the score of a film, the principal purpose of which is to reference a particular theme or character within the diagesis. Leitmotifs in film music are not only used to reinforce a character's presence onscreen but can be used to refer to that character when he/she is absence. In *Cinejack*, *Leitmotif* modules can be set up in conjunction with *Keycheck*, to listen for a particular phrase and use this to trigger an action in the form of a cut, fade, camera change or movement change. The original motif can either be entered note by note, or played in through a 'listen' function.

Changing the tempo of the musical score can radically affect the viewer's perception of the pace of the film and in conjunction with editing can convey a variety of effects. Rhythm and meter can often achieve similar results: for example doubling the speed of the percussion from ½ notes to ¼ notes. The *Tempo* module takes incoming MIDI notes and, after filtering out chords (using Max's *thresh* object) counts notes per bar. This number is used to increase or decrease cutting speed through comparison to a running average. Its output is a message to cut, which can either be passed to straight to a *Cut/Fade*module (to cut to a video track at random) or used in parallel with *Content*, *Tilt* or *Intensity* to cut to specific tracks.

Rises and falls in the relative volume of parts of a soundtrack or a soundtrack's volume in comparison with dialogue and sound effects can be used to intensify or highlight particular themes and are extremely effective in generating tension. Complete silence in the score can be used in contrast with scored sections of the film to lend an immediacy and heightened sense of realism to a scene. The *Intensity* module is used to monitor average amplitude or MIDI velocity (the force with which an instrument is played). Taking inputs from the outlet of *MIDIListener* or *AudioListener*, it attempts to detect crescendos, by comparing each note with its predecessors and can apply the result to devices for dramatic intensity such as slow zooms, or decreases in cutting speed.

Other modules are used to manipulate the video output of *CinejackPlayer*. Content allows filtering of cutting and framing decisions according to the content of each video track. Video events can be tagged in a text document, either and used to specify which video tracks are available to the cutting scheme and keep a constantly updated summary of their content in terms of subject and shot type. *Content* keeps track of its position in the film by receiving its frame numbers via an OSC connection. *Zoom* enables video tracks to be reframed on the fly, by cropping the video image. Used in conjunction with *Content*, *Frame* can look up which type of shot a video track is currently displaying and reframe as a closer shot, for instance rendering a wide shot as a closeup, while taking into account placement of the subject.

Tilt takes inputs from *Keycheck* and, based on the number of accidentals, applies a value to the rotation of the image, either by sending a midi control value directly to *Cinejackplayer*. The result is a Dutch tilt, the angle of which is updated at each cut. For smaller movements, *Shake* takes discord messsages from *Keycheck* or amplitude messages from either listening module and applies them to simulated camera movements. These can be set according to speed and range to either manifest as tiny drifts or pronounced camera shake.

Movement is a module for high-level control of the interaction scheme. One of the few manually controlled modules, it allows a complete new configuration to be loaded onto all available modules with a single keystroke or using a MIDI controller. This can be used either for different pieces of music within a single performance, or for changes of mood or pace within different parts of a film.

We decided to focus on using pre-rendered video to provide the output to *Cinejack*, as this would allow the greatest possible range of visual content to be use. By using multiple video tracks in parallel and by utilizing the high resolutions of 1080p HD video, we would be able to not only cut between video tracks in real time but by zooming and panning across the image, reframe shots. This approach was initially hindered by MAX/MSP/Jitter's relatively poor video performance: consequently we decided to build a custom video player which could be controlled from MAX via MIDI, in order to overcome video performance issues.

To this end, we developed a standalone output module to handle video output. *CinejackPlayer* is a high-performance, multi-track MIDI-controllable video player. Using the gains in performance afforded by SFML (Simple and Fast Multimedia Library), *Cinejack* supports up to 4 channels of full HD1080p video on a mid-range PC laptop (2.4ghz, 8Gb RAM, 1Gb Video Memory). Realtime video transformations, including mixing, panning, zooming and rotation can be accomplished by configurable MIDI controls while an OSC connection to the main *Cinejack* system is used to read and set frame numbers. An easy-to-read text configuration file can be used to quickly preload video tracks and control configurations. *CinejackPlayer* can

also be connected directly to a MIDI controller and used as a simple video mixer.

Cinejack also contains a number of accessory modules including a configurable audio player to handle synchronised sound in line with the video tracks, an editor to prepare metadata for *Content* modules and a customizable manual control module so that some or all of *Cinejack*'s functions can be overriden manually using hardware MIDI controllers. This module is intended mainly for use in the case of technical failure of an audio input, however, it can also be used to trial video transformations before assigning automatic control through modules.

CONFIGURATIONS

As discussed, after the initial scoping project, *Cinejack* was developed through two very different performance projects, which were intended to explore different use cases. The two examples below discuss the development of separate configurations for *Golden Shroud* and *Orientation* and demonstrate two different approaches to using *Cinejack*.

Setup 1: *Golden Shroud*

The first setup was developed in conjunction with an established singer/songwriter (referred to here as Ellen) who writes and records music involving elements of British folk music and Doom metal. Our first step in working with Ellen was to informally discuss how she might imagine visuals supporting her work and try to elicit from her her own views on the aesthetic content and purpose of her work. Next, we established a schedule of live shows in which to demonstrate and test the work. Importantly these were decided on by Ellen, an approach which both guaranteed the authenticity of the setting and minimized the cost to her of her involvement in the project. The concert would be showcases for Ellen's latest album, *Golden Shroud* and would feature Ellen and a band consisting of a drummer and two backing singers. The performances ranged from small bar gigs to a progressive rock festival and presented a wide variety of unpredictable and uncontrolled settings.

Over several months we developed bespoke video content and multiple iterations of the *Cinejack* system simultaneously, attempting throughout to integrate the development into the band's practice with as little disruption as possible. Video footage and animation was produced by the development team in response to an initial set of ideas provided by Ellen. Visual material responding to these ideas was presented to Ellen and her band along with each iteration of the system in her own rehearsal sessions.

During rehearsal sessions, a member of the development team would effectively rehearse alongside the band, adjusting and tuning the responses of the system, developing modules and reconfiguring mappings, with the result projected on the rehearsal room wall. In this way, we effectively borrowed the improvisational techniques of simultaneous composition and rehearsal practiced by the musicians themselves. After each rehearsal, the team would reconvene to firm up the design and implement new features.

We found in particular that these sessions were vital in identifying and solving technical issues (such as dealing with rapid changes in audio inputs as musicians adjusted their instruments) but also helped us understand how Ellen planned and prepared her shows. Being able to compare our own aesthetic decisions with those of others led to a broader understanding of the range of approaches that other musicians might use in composing their own visuals. On a number of occasions, we and Ellen arrived at opposite approaches to particular problems, especially in terms of conceptual approaches to audio/visual mappings. For example: our own use of visuals tended to directly reflect the energy of the band onstage. By contrast, Ellen proposed using visual imagery to establish the overall narrative arc of her stage show, without necessarily following directly the action directly, stating that there were *'peaks and troughs'* in the performance. Key moments were identified as landmarks within the piece, for example *'We're starting it three part a cappella. So we need something that's going to burst'*.

In many cases the ideas she suggested were quite specific, for example runes and symbols were suggested as recurring motifs. In general, however, the narrative element to her visual ideas were far less explicit than in conventional cinema, being more concerned with establishing atmosphere and evoking certain emotional states than using characters and plot. This led us to develop her version of the system to function as a background or virtual set, rather than fully narrative cinema piece, with the audience's attention carefully balanced between visuals and performers.

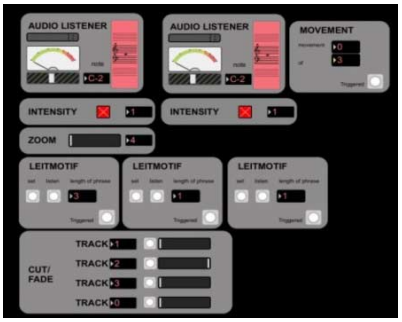

Figure 3. A screenshot of Cinejack's UI set up for *Golden Shroud*. 2 *Audio Listeners* connected to microphones pass on note and amplitude information to *Leitmotif* and *Intensity* modules to affect a *Cut/Fade* module.

Ellen's *Cinejack* configuration (see figure 3.) is controlled by 2 audio inputs: a vocal microphone and an instrument microphone attached to a drum kit, monitored by a separate *AudioListeners*. These are variously passed through

Leitmotif and *Intensity* modules to a *Cut/Fade* module. As discussed, Ellen's approach to visuals was to use them to support the broad narrative arc of her show. This setup was designed to facilitate this by allowing her voice to control various changes in a visual scheme which would change from movement to movement.
Several *Leitmotif* objects were used to respond to vocal notes and phrases to generate effects such as sudden cuts to bursts of flame. An *Intensity* object mapped to cut objects caused the video tracks to crossfade in particularly loud passages in the track. This was used to partially fade in video tracks, creating movement across the visuals. A *Zoom* module was also mapped directly to the amplitude of the signal from the drums. This caused the whole video image to pulse rhythmically during various points in the show. 3 different phases in the performance, related to the 3 musical pieces played, were reflected by changes in the video scheme, triggered manually using a *Movement* module linked to a korg MIDI controller.

Setup 2: Orientation

Golden Shroud allowed us to develop and test *Cinejack* in real concerts as a tool for augmenting existing performances, through a dialog with other practitioners. Allowing another musician to set the agenda entirely in terms of her own practice meant that we could guarantee the appropriateness of our design decisions, however due to the pre-determined requirements of Ellen's practice and the lack of space for experimentation, we felt that we had not fully developed *Cinejack*'s narrative potential in a way that we could confidently generalize for a wide range of musicians. To this end, we decided to set up a collaborative project whereby a musician would compose music to a visual scheme established by us and would specify the connections and interactions the system should have. We anticipated that this would form a bridge between two creative practices, while still providing an authentic context in which to work.

To provide a test-bed set of narrative visuals we embarked upon a full-scale film production project: writing and directing a multi-channel short horror film entitled *Orientation*. Shot on location using a professional cast and crew, the plot of this piece followed Rachel, a woman confronted by mysterious figures and obstacles as she tries to find her way out of a seemingly endless series of tunnels. A genre film of this type seemed particularly appropriate as a test-bed, as it provided a readymade set of particularly powerful cinematic and musical conventions while also allowing experimentation. Shot in 1080p HD, each scene of *Orientation* consisted of multiple video tracks comprising a range of camera angles, designed to be reframed and edited by *Cinejack* in real time.
Like the original scoping project *OverWatch*, *Orientation* was designed to be performed in a live 'Silent Cinema' setting with different edits potentially provided by improvised scores from a number of musicians. The first of these, (referred to here as Jane), was a classically trained pianist with extensive experience of live improvisation. We invited her to prepare a score which responded to the narrative content of the movie and which left space for improvisation on her part in a live performance. This score, played through *Cinejack,* would re-direct the movie, changing the cinematography in terms of editing, framing and camera movements.

Once again, further development of *Cinejack* was 'performed' in rehearsal with Jane as she played over the film. By recording and then later revisiting these sessions, we were able to ask specific questions about her decisions during the composition process, such as her choice of leitmotifs, how she felt the music connected with the visuals and how she considered the narrative arc of the movie in musical terms. Through discussion and extensive rehearsal and development sessions, we arrived at a set of mappings and controls which Jane felt were legible and appropriate to both the score and the film. We agreed that the pace of the film should be determined by the tempo of the score. She composed leitmotifs to signify certain events in the film and to refer to characters and requested that these trigger cuts to those characters. This led to us developing the *Content* metadata system to keep track of characters across the various video tracks, so that the system could cut to them whenever necessary.

Several new challenges arose in Jane's subtle approach to music-making. In contrast to our own and Ellen's approaches, she intended to include passages where there was no music at all. Reflecting on the score, she commented, *'I wanted silence and simplicity to be part of it…making sure that it doesn't completely take over'.* This meant that we had to prepare *Cut/Fade* modules to enter an idle phase, during pauses in Jane's playing. The complexity of the film itself, led to a number of challenges. Not least of these was helping Jane navigate the multiple camera angles while playing. Our initial solution to this was to provide cues on the monitor of the computer running *Cinejack*, however, Jane declared that this would split her attention between playing and watching the film on the screen. This led us to devise an optional cueing system in *Content* whereby the entrance of a new character into the action on any video track could trigger a cut to that track.

The setup in figure 4, taken from the performance configuration for Orientation comprises an elaborate interpretive scheme to control cutting and shot selection. Input in this case is provided through a MIDI connection to a digital piano. A *Content* module is used to keep track of 9 different subjects in the film, from characters to important objects. This module is preloaded with a video annotation document, constantly keeping *Cinejack* updated with a list of objects and characters visible on each video track. The module can then select which video track to cut to each time a *leitmotif* module referring to a particular character

was triggered. In this setup, *leitmotif* modules are configured to listen to several different phrases which Jane used to variously refer to the main protagonist, a recurring image of a gas-mask and a mystery figure in the drama who appeared at several points. When these phrases are played, *Cinejack* automatically cuts to a video track depicting the relevant subject.

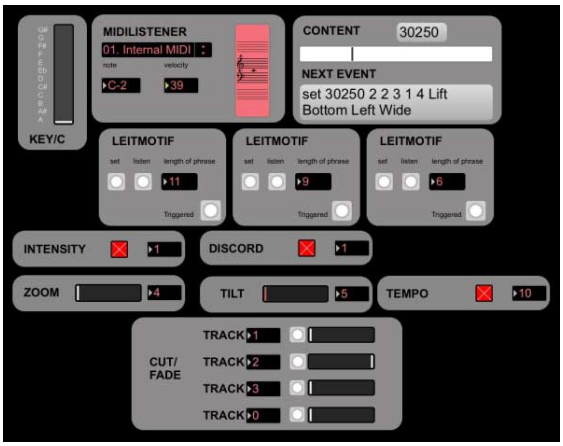

Figure 4. Screenshot of the performance setup for *Orientation*. A single *MIDIListener* monitors the digital piano. Leitmotifs, notes outside the key and crescendos are mapped respectively via *Leitmotif*, *Intensity* and *Discord* modules to *Cut*, *Tilt* and *Zoom* modules.

According to Jane's requests, cutting speed throughout the film is determined by a Tempo module, monitoring the pace of the score. Discordant notes, triggered from a *Keycheck* module are linked to a tilt module, generating Dutch tilts in particularly nightmarish sequences, while crescendos are mapped via a *zoom* object, to slow tightening of the frame, emphasizing a sense of tension.

DISCUSSION AND CONCLUSIONS

In the development of *Cinejack* the combination of multiple perspectives and testing in authentic and unpredictable situations proved highly effective in gaining a deep understanding of the design space. This approach yielded a wealth of useful information ranging from specifics of how to appropriately map particular instruments to cinematic effects, to technical experience in maintaining useful levels of audio input in the cluttered sonic environment of a rehearsal room. More importantly, it has stimulated a far deeper understanding of other musicians' performance practices in context, in a way which we believe will ensure the appropriateness and usefulness of the final set of tools. The current version of *Cinejack*, recently released in open-source form at XXXX meets our initial design constraints in being simple to integrate into musicians' own practices and - according to our participants - in adding significant value to their stage shows. The effectiveness of *Cinejack* as a tool for controlling visuals was confirmed by our participants. Regarding the visual scheme designed for *Golden Shroud,* Ellen stated,

"I think it adds to not only the mood but the continuation of the pieces. Because, you know, you want it to feel like every track is part of a joint venture to one purpose".

Our method of appropriating the venues and compositional techniques of our target users was highly productive as it allowed us to establish a rapid development cycle, where features could be requested, implemented and evaluated in the same rehearsal sessions. Issues of stability were quickly eliminated through setup and play in multiple rehearsal rooms, stages, theatres and (in the case of the festival) a remote hillside. By working entirely in authentic settings, we were able to both accelerate the information gathering and testing phases of the research and also achieve a better understanding of the exact context of the design.

The development process was challenging for the researchers involved, involving not only leaving the relative comfort and convenience of the laboratory but also structuring the development process entirely around our participants' schedules, however an unanticipated benefit of our long-term engagement with the participant musicians was the growing emotional investment of the development team in the success of the test-bed shows. While presenting new footage and new iterations of the system to the musicians, the researchers felt exactly the same nervousness and stage fright we would expect to feel playing for the first time with unfamiliar musicians. We found that the added pressure of 'performing' each iteration in rehearsal led, on the one hand, to a deeper understanding of each musician's practice in its proper context and, on the other hand, to a more authentic method of testing the system, highlighting bugs or design flaws that might otherwise have gone unnoticed.

In structuring the research around our participants' practices, we were careful to minimize the risk to their work. Our participants were able to confirm that not only had we not significantly disrupted their work but the research had been directly beneficial to them. Jane stated that the novelty of the format and the risk of technical failure in rehearsals had initially caused her anxiety.
"I knew how many things could go wrong and how all these different systems had to work with the visuals and the music and editing as well".
However, although the project had been daunting, she found it had allowed her a new perspective on her own practice and to reflect on her own decisions in a way that was not only enlightening for us but also stimulating for her. On participating in the project she stated,
"For me it was something really positive and really amazing to be a part of and really developed me as a performer and a musician".

It is already becoming common practice for researchers in Human Computer Interaction to work alongside practitioners to better understand their domain. We suggest

that our approach - using multi-disciplinary teams to develop and use designs firsthand in structured collaborative dialog with their target users - offers a way to quickly and efficiently develop tools, the appropriateness and usability of which is ensured by the authenticity of their development.

The core approach of appropriating users' own musical instruments has been particularly successful in terms of overcoming the traditional learning curve associated with learning a new mode of interaction. Using their own instruments, meant that even in early rehearsals, participants were immediately able to 'play' the system without feeling overwhelmed by the additional set of controls they were commanding and were satisfied with the resulting effect of the visuals. In both projects, the musicians reported being able to perform without feeling distracted and feeling that the visuals were actively enhancing the show. Reflecting on the *Golden Shroud* concerts, Ellen commented,
'It was like another band member, it was like somebody joining in'.

We believe that the design strategy described here, although logistically demanding, could be applied to other domains, including those outside the arts. The design requirements provided by our participant musicians have much in common with other cases where expert practice is involved. Numerous other disciplines involve highly-skilled user groups, working in difficult environments with severe logistical or structural constraints. By capitalizing on users' own skills and the affordances of tools with which they are already have a highly developed relationship, designers might develop new tools which not only extend the user's ability in new ways but can be easily integrated into existing working practices.

REFERENCES

1. Makela, M. The Practice of Live Cinema. *Media Space Journal* (2008)
2. Kalinak, K. *Settling the Score, music and the classical Hollywood film*. University of Wisconsin Press (1992)
3. Bongers, B. and Harris, Y. A Structured Instrument Design Approach: The Video-Organ, *In Proc. New Interfaces for Musical Expression (NIME)*, (2002), 18–23
4. Bevilacqua, F., Rasamimanana, N., Flety, E., Lemouton, S., and Baschet, F. The augmented violn project: research, composition and performance report. *In Proc. New Interfaces for Musical Expression(NIME)*, (2006)
5. Overholt, D. The Overtone Fiddle: An Actuated Acoustic Instrument. *In Proc. New Interfaces for Musical Expression (NIME)*, (2011)
6. Newton, D. and Marshall, M. Examining How Musicians Create Augmented Musical Instruments. *In Proc. New Interfaces for Musical Expression (NIME),* (2011)
7. Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M. The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. *In Proc. 1st international Conference on Tangible and Embedded interaction (TEI 2007*. ACM Press (2007)
8. Magnusson, T., & Mendieta, E. H. (2007, June). The acoustic, the digital and the body: A survey on musical instruments. In *Proc. New Interfaces for Musical Expression (NIME).* (2007) 94-99
9. Stam, R., Burgoyne, R. and Flitterman-Lewis, S. *New Vocabularies in Film Semiotics*. Routledge, London (1992)
10. Arijon, D. *Grammar of the Film Language*. Hastings House Publishers (1976)
11. Katz, S. *Film Directing: Shot by Shot*. Michael Wiese Productions, (1991)
12. Brown, R. *Overtones and Undertones, Reading Film Music*. University of California Press, (1994)
13. Manovich, L., Kratky, A. *Soft cinema: navigating the database*, MIT Press, (2005)
14. http://www.cfccreates.com/what_we_do/cfc_media_lab/interactive_narrative_feature_program/index.php
15. Lino, C., Christie, M., Lamarche, F., Schofield, G., and Olivier, P. A Real-time Cinematography System for Interactive 3D Environments. *Proc. Eurographics/ ACM SIGGRAPH Symposium on Computer Animation* (2010), 139-148
16. Thompson, W,F. Graham, P. and Russo,F. Seeing Music Performance: Visual Influences on Perception and Experience. *Semiotica*, (2005), 156
17. Hook, J., Green, D. and Olivier, P. A short film about VJs: Using documentary film to engage performers in design. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI'09)*, (2009)
18. Reeves, S., Benford, S., O'Malley, C. and Fraser, M. Designing the spectator experience In *Proc.CHI 2005* ACM Press, (2005)..
19. Sheridan, J., Dix, A., Lock, S. and Bayliss, A. Understanding Interaction in Ubiquitous Guerrilla Performances in Playful Arenas. *In Proc. HCI*, Springer, (2004), 3–18
20. Buxton, W. Artists and the Art of the Luthier. Computer Graphics: The SIGGRAPH Quarterly, 31(1), (1997).10-12.