Creative Collaboration Between Audiences and Musicians in *Flock*

Jason Freeman, Mark Godfrey
Georgia Tech Center for Music Technology
840 McMillan ST
Atlanta, Georgia USA 30332-0456

J. Freeman and M. Godfrey. "Creative Collaboration Between Audiences and Musicians in Flock." Digital Creativity, Vol. 21, No. 2, 2010.

ABSTRACT

Flock is a full-evening work for saxophone quartet, dancers, audience participation, video, and electronic sound. A computer vision system tracks the location of participants as they move around an open performance space, and custom software then uses that data to generate music notation, video animation, and electronic sound. In this paper, we discuss the project's connection to previous work in large-audience creativity and multi-player gaming, and we outline the conception, design, and implementation of its computer vision system, real-time notation architecture, and collaborative structure. We evaluate the project's success as measured by audience surveys from performances at the Adrienne Arsht Center for the Performing Arts in Miami and the 01SJ Festival in San Jose, and we outline revisions to the work based on that feedback.

KEYWORDS

music, computer vision, positioning, real-time music notation, audience participation, saxophone, dance

1. INTRODUCTION

Social networking services, consumer-oriented production software, and online distribution mechanisms have made our experiences with recorded music increasingly interactive, creative, and collaborative. We curate our own playlists, share them online, and seek music recommendations based on our listening habits. In spite of legal obstacles, we remix and mash up the music we love in order to create compelling new works. And using popular new services (e.g. SpliceMusic 2008), we collaboratively create music by drawing from a shared library of sound clips and a tree of remixes.

Yet as recorded music inspires new models for collaborative creativity, the few-to-many model of live concert performance remains largely unchanged. The real-time nature of performance, the large audiences typically in attendance, and the physical and technological constraints of venues make it difficult to incorporate the same digital tools into live performances that have been so successful with recorded media online.

In *Flock* (2007), a full-evening work for saxophone quartet, dancers, audience participation, electronic sound, and video, we attempt to create a collaborative environment that links together the creative activities of composers, performers, and listeners in live performance. Using novel computer vision and real-time notation systems, we delay content creation until the moment of each performance, so that the music can reflect the activities of each show's performers and audience members. Music notation, electronic sound, and video animation are all generated in real time based on the locations of musicians, dancers, and audience members as they stand up, move around, and interact with each other.

Flock aims to make its audiences feel like participants who shape the performance, not spectators who merely watch it. We want audiences to feel connected to the musicians, the music, and to each other, to discover new ways to be creative, and to realize that each performance is unique, in part, because of their contributions to it.

2. BACKGROUND AND PREVIOUS WORK

2.1. Audience Participation

Flock follows in the tradition of musical works that incorporate audience participation. In many such works, audience members become performers, creating some or even all of the music. For example, in *Moths* (Hasse 1986), the audience whistles as directed by a conductor and a graphical score to perform the piece. During *La symphonie du millénaire*, an outdoor performance event in Montreal, 2000 audience members rang handheld bells at designated times (Chénard 2000). And many Fluxus scores specify or imply more open-ended audience participation, as with Tomas Schmit's *Sanitas No. 35*: "Blank sheets are handed to the audience without any explanations. 5 minutes waiting" (Schmidt 1962).

Other musical works invite the audience to shape the music performed by live musicians rather than creating sounds that are part of the performance. In *No Clergy* (Baird 2005), computer software stochastically generates notation for each performer based on votes cast by audience members on laptop computers. McAllister et al (2004) developed a performance environment in which audience members draw notation on a PDA's touch screen for musicians to play. And

Wulfson, Barrett, and Winter (2007) created *LiveScore*, in which gallery visitors adjust knobs on physical controllers to modify the parameters of a stochastic music notation algorithm.

Flock is most closely related to this second category of works. Our past experience had convinced us that many audience members were comfortable creating their own sounds via the anonymity of the Internet (Freeman et al 2005), but at a live public event, their self-consciousness inhibited their creativity. Additionally, this paradigm offered a live-performance parallel to the ways in which listeners creatively engage with recorded music: rather than replacing professional musicians altogether in performance, the audience creatively reshapes their music.

2.2. Multi-player Gaming

Flock is not a game, but it does include implicit competitive elements. On their own, audience members play only a small role in the performance, but if they can convince other participants to follow their lead, they can dramatically influence the show.

Flock also draws from techniques developed for multiplayer games staged in physical venues. Some of these invite seated audience members to play collaboratively by holding up a colored paddle (Carpenter and Carpenter 1999) or shifting in their seats (Maynes-Aminzade, Pausch, and Seitz 2002), while others take place in a mixed reality space combining location-aware devices with physical game play in a venue, neighborhood, or city (e.g. Flintham et al 2003). And multiplayer games such as *Sonic Tug of War* (Feldmeier and Paradiso 2007) focus on musical goals.

2.3 Previous Work

Flock builds upon previous work by one of the authors. In *Glimmer* (Freeman 2008a), a composition for chamber orchestra and audience participation, each audience member is given a battery-operated light stick that he or she waves back and forth over the course of the piece to influence the music (Figure 1). The audience is divided into groups, and a computer vision system analyzes light-stick activity in each group; software then uses the analyzed data to generate musical instructions for corresponding groups of orchestral musicians. Each musician has a computer-controlled, multi-colored light mounted on his or her music stand to display the notes, dynamics, and articulations generated by the real-time algorithm.

[Figure 1. Audience members in Jason Freeman's *Glimmer* wave light sticks back and forth to shape the music played by a chamber orchestra.]

While performances of *Glimmer* have been fun, engaging, and at times musically satisfying, the work has several significant problems that we sought to address in *Flock*. *Glimmer*'s large scale — 600 audience members and 25 musicians — makes it difficult for every participant to play a significant role. So with *Flock*, we downsized to 100 audience members and a quartet of musicians. With *Glimmer*, the traditional concert hall layout emphasizes the separation of musicians and audience and can even make it difficult for audience members to see each other. So we staged *Flock* in a blackbox space with flexible seating. Most of the orchestral musicians performing *Glimmer* have been uncomfortable improvising, so they follow their instructions exactly and their creative role becomes limited. So we scored *Flock* for jazz musicians with expertise in improvisation. Finally, the realities of orchestral program conventions have limited the length of *Glimmer* to just 10 minutes — hardly enough time for a meaningful collaboration to develop. So we designed *Flock* to last a full evening.

3. DESIGN AND IMPLEMENTATION

[Figure 2. Conceptual illustration of the primary structural data flows in *Flock*. Solid lines indicate digital data links between people and software algorithms, while dotted lines indicate direct perceptual links among participants.]

Flock's architecture creates real-time connections (Figure 2) linking participants during the performance. A computer-vision-based positioning system determines the locations of audience members, saxophonists, and dancers as they move around the open performance space. This location data drives software algorithms that generate music notation, electronic sound, and video animation. The saxophonists interpret the notation to perform the music. All of the participants then respond to the saxophone music, the electronic sound, the video animation, and each other, changing their positions and beginning another iteration of the interaction.

3.1. Computer Vision Positioning System

In order to facilitate creative collaboration with a large, live audience, *Flock* must track the locations of participants as they move around the performance space. We developed a computer-vision-based multi-target positioning system to track location data. Although we investigated other tracking methods, such as triangulation via 802.11 signal strength, RFID, inertial navigation, and GPS, computer vision was the only solution which met all of our needs: adequate precision over a large area, reasonably fast frame rates, indoor functionality, easy scalability to a large number of targets, fast installation and calibration time, and low cost.

Flock follows the example of numerous interactive artistic and musical works that have also employed computer vision to facilitate expressive interactions with technology by tracking the positions of participants. For example, *Boundary Functions* (Snibbe 1998) tracks participants' locations and dynamically projects a set of lines onto the floor that divide participants from one another. In *Bodymover* (Sauter 2000), participants collaboratively compose an audio-visual experience based on their camera-tracked interactions with virtual sound objects. *Messa Di Voce* (Levin and Lieberman 2004) uses performers' locations, along with live audio analysis, to drive a video animation, creating a synesthetically augmented reality. Shaun Moon's *Bach Blocks* (Do and Gross 2007) focuses on the locations of objects rather than people: users create music by arranging blocks on a grid, and their position and color is used to generate musical motives. Similarly, the *reacTable* (Jordà et al 2007) is an expressive tabletop interface for live music performance that uses computer-vision-driven pattern recognition to identify objects on its surface; each object represents a sound function whose parameters are manipulated by a user moving the object.

Flock's positioning system tracks the locations of one hundred audience members, four dancers, and four saxophonists within the space; each person is represented as a single point on the two-dimensional plane, seen from a birds-eye view and parallel to the venue's floor. The system identifies each saxophonist with a unique label but does not distinguish among other points.

The entire performance area is captured with a single ceiling-mounted firewire camera fitted with a fisheye lens. Our software, implemented in Cycling '74's Jitter with the cv.jit library (Pelletier 2008) and our own custom objects, first performs lens correction to account for most of the fisheye distortion. This approach has proven more robust than using multiple cameras with minimal lens distortion; our image skew and stitching algorithms increased setup time, reduced frame rates, and created small image boundary errors.

Each audience member and dancer wears a baseball hat with a white LED sphere to facilitate tracking. Camera iris, shutter, and exposure settings, combined with a luminosity threshold filter, isolate the LED lights on the image, and connected components analysis detects blobs and their centroids. Though we had not originally wanted the audience to wear lighted hats, this improves reliability under variable lighting conditions with minimal calibration, eliminating problems caused by shadows and clumps of people. The novelty of the hats also helps the audience get excited about participating. And we preferred the white LEDs to an infrared solution; in our experimentation, we needed far more and far brighter IR targets in order to track at the same distances and camera skews.

[Figure 3: Saxophonists wear baseball hats with uniquely-colored LED spheres to facilitate tracking via chromakey and a particle filter.]

Each saxophonist wears a hat (Figure 3) with a unique LED color in its sphere. A chromakey produces a "distance image" to represent the similarity of each source pixel to each target color. Pattern-based identification was not practical because of the extreme skew as targets moved to the edge of the image.

[Figure 4: a): This raw, unanalyzed video frame includes a target — the light on the performer's head — and a variety of other elements that should not be tracked. b) Chromakey filtering eliminates most of the unwanted elements, but two blobs similar in color to the target remain in the image. c) The particle filter has been trained to lock on the desired target, and it follows that target from frame to frame. The unwanted elements are ignored.]

At this stage, lighting conditions of the venue as well as the unpredictable color makeup of the audience can introduce considerable noise into the chromakeyed image, rendering simple blob detection inadequate for tracking our targets. So we use a particle filter to track each saxophonist. The particle filter employs importance sampling from a set of particles moving stochastically across an image from frame to frame: the closer a particle to its target, the more likely it is to survive into the next frame. In this way, samples that are statistically insignificant die out, moving the entire system of particles towards the saxophonist. The algorithm reports the location as the weighted centroid of all particles. Based on the approach developed by Nummiaro, Koller-Meier, and Van Gool (2003), our custom-built particle filter (Figure 4) uses the output of the chromakey as the probability distribution from which to sample, giving us control over the noise tolerance by adjusting the chromakey parameters.

The entire algorithm operates at approximately 6 fps on our 2.4 GHz MacBook Pro. While this is generally adequate for our needs, analyzed motion can become jerky when targets move extremely quickly; we compensate with a smoothing filter in the video animation and some connected line segments in the music notation.

3.1.1. Flock Vision Toolkit

We have released the core components of our computer vision system, along with other tools developed for earlier experimental setups, as an open-source software package. The Flock Vision Toolkit, available at http://www.jasonfreeman.net/flock/, includes computer vision objects and utilities for Cycling '74's Jitter environment.

The toolkit includes a particle filter tracking object; a skew correction object to compensate for camera perspective; a stitching object to combine images in multi-camera setups; and a barrel/pin correction object to compensate for fisheye lens distortion.

3.2. Real-Time Music Notation

In most live computer music performance environments, software algorithms either directly output digital audio, or they output control data for other music-making devices (such as MIDI synthesizers or mechanical automatons). Human musicians either improvise or play from a printed score. They respond to the computer-generated sound they hear, but they do not directly interpret its output.

This paradigm works well in many performance situations, creating distinct musical voices from human and machine musicians. But with works such as *Flock* that link audience participation to musical output, it has a significant drawback: the audience's input generates electronic sound but only indirectly affects the human musicians, and so the connection between the audience and musicians is diminished (Figure 5).

[Figure 5. a) Simplified data flow in an interactive system in which the audience influences real-time generation of electronic sound while musicians play from a predetermined score. b) Simplified data flow in an interactive system in which the audience influences real-time generation of music notation.]

Real-time music notation systems, in which visual scores are dynamically generated and displayed, provide a powerful alternative to more deeply connect musicians and audiences. Audience input data drives a software algorithm that generates the music notation in real time to direct the musicians' performance, bringing all participants into a tightly linked interaction.

Real-time notation systems draw from two important musical traditions of the last half-century. In computer-assisted composition, non-real-time algorithms generate musical scores, event lists, or score fragments that composers assemble into instrumental scores; examples include the *Illiac Suite* (Hiller and Isaacson 1959), Experiments in Musical Intelligence (Cope 1996), the Free Stochastic Music Program (Xenakis 1992), and the Patchwork and OpenMusic software packages (Assayag et al 1999). Meanwhile, open-form compositions such as *Available Forms I* (Brown 1961) and *Klavierstuck XI* (Stockhausen 1957), in which the performers determine the order of musical fragments, established precedents for the transformation of score materials during live performance.

As real-time notation rendering and distributed display have become more practical, more composers have begun to incorporate real-time notation systems into their work. Recent examples include Nick Didkovsky's *Zero Waste* (2004), in which a pianist sight-reads real-time notation as her errors are transcribed to create subsequent measures of the piece, and David Kim-Boyle's *Valses and Etudes* (2005), in which excerpts of famous piano scores are stochastically manipulated to create real-time notation. Art Clay's *GoingPublik* (Gutknecht, Clay, and Frey 2005), for trombone trio, is of particular relevance to *Flock*; GPS units and 3D compass sensors detect the position and movement of each musician to generate the real-time notation. For a more complete discussion of real-time music notation, see Freeman (2008b).

In *Flock*, real-time music notation is generated for each saxophonist based on the data from the positioning system. Our Java software, built using the Java Music Specification Language

(Didkovsky and Burk 2001), renders images and sends them wirelessly to PocketPC devices mounted on each player's instrument using marching-band lyres (Figure 3). Our PocketPC software, also written in Java, decompresses and displays the images, maintains time sync, and renders a scrolling measure-position bar.

3.2.1. Notation Styles

Our software uses a variety of approaches to render music notation. In rhythmically sparse sections, it displays notes on a conventional music staff (Figure 6). Though pitches are exact, time is proportional; horizontal lines extend from solid noteheads to indicate the approximate duration of each note.

Rhythmically denser music is difficult to cleanly render and sight-read on a staff, so for these passages, we switch to a graphical view (Figure 7) that emphasizes gestural contours. Notes appear on a time-pitch grid; successive octave registers and a handful of notes are labeled with exact pitches. Diagonal lines often connect nearby noteheads.

[Figure 6. Staff-based music notation. The musician plays the Db, which is rendered in green. The remaining notes, rendered in pink, show what the other saxophonists play.]

[Figure 7. Graphical music notation. The musician plays the lower-left contour, which is rendered in green. The remaining gestures, rendered in pink, show what the other saxophonists play.]

Both the staff-based and graphical views display only a single measure of music at a time, due to limited screen size (4 inches diagonal with 320 x 240 pixel resolution), as well as the need to minimize the delay between the music's rendering and its performance. But musicians need to read music ahead — especially when sight-reading — so they can prepare for what comes next. We employ two strategies to accommodate this need. In the first, notation is updated frequently during each measure, as each new frame of position data is calculated. In this manner, the notation changes gradually over the course of each measure, rather than suddenly at the beginning of each measure. In the second strategy, measures are initially frozen. But as the measure is played and the position bar moves from left to right, the played portion of the measure begins to update with content for the next measure, using the same technique of frequent, gradual updates as in the first strategy. We use both strategies in the performance, depending on the density of musical material, its rate of change, and the importance of shorter latency times.

The size and brightness of noteheads, and of lines connecting noteheads, indicate their dynamic. Articulations are drawn above noteheads as in conventional notation. Each musician's own notes are rendered in green; the other players' music is rendered in a transparent pink. Textual cues instruct musicians how to move and interact on stage and indicate the progression through the structure of the piece.

3.2.2. Data Mapping

In some sections of the performance, each musician generates his own notation via a simple mapping of his two-dimensional birds-eye location in the space: his x position maps to the measure position of the note onset, and his y position maps to its pitch (Figure 8a). Via a similar mapping, dancers and audience members generate notes for the musician closest to them.

In other sections of the piece, dancers and audience members generate notation based on their locations relative to each musician (Figure 8b). Each saxophonist is the center of his own polar coordinate system, and any person within a maximum radius generates a note; the distance between them determines the note's pitch, while the angle determines its measure position. In this manner, as more participants draw closer to a musician, his notation includes more events per measure.

[Figure 8: Mapping strategies in *Flock*, shown from a birds-eye view of the performance space. a) Each participant generates an event based on her Cartesian position in the space; x maps to measure position and y maps to quantized pitch. b) Each saxophonist (S) is the center of his own polar coordinate system. The distance (r) of participants (P) maps to the quantized pitch of events, and their relative angle (θ) maps to the measure position of those events.]

In many sections of the performance, notation is also generated from a history of data going back several seconds; participants essentially leave trails on the notation as they move. This helps to create more continuous gestures in the notation, build up more complex and interesting musical textures, and facilitate more gradual musical change from measure to measure.

In the sparsest sections of the music, the notation indicates exact pitch and timing information for every note, leaving little room for interpretation. In the densest sections, the notation includes clouds and clusters of notes with less specific pitch and timing information. It is no longer possible for musicians to play each individual event precisely; the notation instead becomes a guide for improvisation, indicating texture, density, register, and contour.

3.3.Electronic Sound

While *Flock*'s focus is on the acoustic sound created by the saxophonists, some sections use electronic sound to underscore the saxophones or to create music while they are offstage. In these sections, each audience member's Cartesian position generates a single synthesized note per measure: x corresponds to measure position and y to quantized pitch.

Our software, written in Cycling '74's Max/MSP, relies on CNMAT's spectral manipulation objects (Zbyszynski, Wright, and Campion 2007) and on the RTcmix external (Garton and Topper 1997) for physical models of plucked string and struck percussion instruments. Small variations in audience positions modify the timbres, dynamics, and envelopes of the sounds, and each note is diffused over an eight-speaker system according to the person's location.

Electronic sound also plays a prominent role in the interactive environment that the audience encounters as they enter the performance space. This component of *Flock*, titled *Covey*, is inspired by the rich timbre and fluid gestural control of the Aeolian harp, an instrument "played" by the wind. Our computer vision system tracks participants' locations and movements, translating them into "wind energy" to play the synthetic harp. Four virtual strings, each tuned to a separate pitch, span the width of the space (Figure 9), and as an audience member approaches a string, it is increasingly excited, causing it to sonically emerge from the ongoing bass rumble. The position along the string at which it is excited emphasizes certain harmonics from the string's timbre, just as a guitar has a brighter tone when its strings are plucked near the bridge. These timbral changes all occur continuously and allow for a flowing, meditative aural landscape, similar to an Indian tambura. The tuning of the strings also changes slowly, the resultant chords subtly introducing the audience to the harmonic context of the entire piece.

[Figure 9: In *Covey*, the first component of *Flock*, a birds-eye view of the performance space is divided into virtual strings (horizontal lines) and harmonic partials (vertical lines).]

As audience members enter the space and interact with *Covey*, ushers encourage them to explore how their position in and movement through the physical space transforms the sound of the virtual harp and affects the complex timbral space, preparing them for their role in the subsequent sections of *Flock*.

3.4. Video Animation

In addition to real-time music notation and electronic sound, *Flock*'s computer software also generates video animation. Since audience members rarely see the actual music notation — they would have to look over a saxophonist's shoulder to do so — the animation provides a stylized interpretation of the same musical data. This animation helps participants to understand the notation and electronic sound generated by the algorithm and their role in shaping it.

For the premiere performances of *Flock* in Miami, we invited visual artist Liubo Borissov to design the real-time video animation. Borissov's video, created in Cycling '74's Jitter with Lua (Wakefield and Smith 2007), was projected onto the four walls of the performance space. It rendered position and notation data onto a three-dimensional grid (Figure 10). Dots represented each participant's location, and notes were drawn as bubbles, trails, and diagonal lines that moved up the vertical axis to reflect measure position and show the connections between musicians and other participants.

[Figure 10: Original video animation created for *Flock* by Liubo Borissov.]

While Borissov's animation was aesthetically appealing, audiences at the Miami performances expressed two main concerns about the video. (See section 4.1 for a complete discussion of audience surveys and feedback.) First, they often found it hard to connect the abstract bubbles and squiggles in the virtual space to the music they heard. Second, they remained curious about the real-time music notation; they wanted to better understand what the musicians saw so that they could appreciate their virtuosic, creative role in the performance.

So for a subsequent set of performances in San Jose, we created a new video animation based more directly on the graphical real-time music notation (Figure 11). We eliminated some aspects of the saxophone notation, such as note name labels, and added other elements, such as color-coding that indicated which music belonging to each saxophonist. For the polar-coordinate mapping mode, the video continues to display data in a Cartesian system but draws arcs which rotate around each saxophonist, radar-style, to indicate the data each musician plays. While this video may be less aesthetically appealing than Borissov's original animation, it communicates the data more clearly to the audience and satisfies their curiosity about the music notation.

[Figure 11: Revised video animation based more directly on the graphical real-time music notation.]

3.5. Structure and Organization

Our biggest challenge in developing *Flock* was not the technical implementation, but rather the structural framework for the participation of the four musicians, four dancers, and one hundred audience members and the progression of that framework over the work's extended duration.

3.5.1. Musical Structure

The pitches notated for the saxophonists and synthesized in the electronic sound are quantized to a pitch set. Over the course of each performance, the software moves through a circular progression of 28 pitch sets (Figure 12); adjacent sets always differ by a single pitch class. With graphical notation, the saxophonists do not limit themselves exclusively to the notes in the pitch set, but the labeled notes indicating important points of arrival always fall within the set.

[Figure 12. The first 8 pitch sets in the 28-set progression. The eighth pitch set is a whole-step transposition of the first. The remainder of the progression is similar in structure to this excerpt, but it moves up and down to different transposition levels. The sets are voiced differently for different instruments and in different contexts; this figure shows a "sparse" voicing for alto saxophone.]

Each performance is divided into different sections broadly defined by their participants — saxophone quartet only; quartet with dancers; quartet, dancers, and audience; or dancers and audience — as well as the coordinate system used to map data — Cartesian or polar. While the order of these sections has varied in different performances of the work, the basic progression they have created remains the same; the performance gradually moves from sparser to denser musical textures, from fewer to more simultaneous participants, from more organized movement to more open-ended activity, and from more fully-notated music to more guided improvisation. Within this larger formal structure, each individual section functions as a small arch, moving up and then down through similar continuums. The peak of the arch for each successive section is higher than the previous. A short coda at the end of the performance returns to sparser musical material and features the dancers and saxophonists on stage alone.

It is difficult to describe the music itself, since the musicians' response to the notation has a tremendous effect on musical style and content. At the premiere performances in Miami, the music ranged from pointilistic bursts and slowly-changing drones to rhythmically dense textures full of sudden register shifts, undulating arpeggios, and multiphonics, calling to mind influences ranging from Terry Riley and Steve Reich to Ornette Coleman and Evan Parker. At performances in San Jose with the Rova Saxophone Quartet, the music reflected the unique improvisational style of that ensemble.

3.5.2. Dancers and Audience Members

[Figure 13. One of the instruction cards given by the dancers to audience members at the Miami performances.]

During test runs of *Flock* before its premiere, we discovered that the audience needed guidance in their participation; otherwise, they became frustrated, paralyzed, and eventually bored by the completely open-ended nature of their participation. To address this, we have recruited dancers to assist in performances of *Flock* and to facilitate the audience's participation.

At each performance, the audience initially remains seated as the saxophonists and dancers perform, introducing the key musical material and conceptual ideas for the piece. Over the remainder of the performance, the dancers invite audience members out of their seats, organize them into groups, and guide them through their movements. At the premiere performances in Miami, the dancers handed out printed instruction cards to audience members (e.g. Figure 13), directing individual audience members to lead their groups according to specific instructions and to then hand off their leadership role to another audience member. In subsequent performances, dancers have been more spontaneous in their guidance of the audience, using physical gestures and whispered instructions in lieu of the pre-printed cards, which had added a layer of logistical complexity to the interaction and had often stifled the pacing of the show. Dancers still do organize the audience according to a general plan developed in advance, but they do so with greater flexibility to respond to the unique creative activities that emerge during each show.

During our test runs, we also found that the creativity of participants was inversely proportional to their number. As the stage became more crowded, it was more difficult for audience members to move around freely, and most of them preferred to wait to follow the actions of others rather than initiate their own creative movements. With fewer participants on stage, each had more freedom to creatively explore her relation to the performance space and to the musicians, and each was better able to sense her contribution to the music and visuals. So in performance, the dancers limit the number of simultaneous people on stage, constantly encouraging new people to participate and others to sit down and temporarily become spectators. In some sections, there are only a handful of audience members on stage; in others, there are as many as thirty.

4. EVALUATION AND DISCUSSION

4.1.Audience Feedback

[Table 1: Results from audience surveys at the Miami and San Jose performances. Audience members indicated the degree to which they agreed (5) or disagreed (1) with each statement. We then calculated these interpolated median responses for each statement.]

With *Flock*, we wanted audiences to feel connected to the musicians and to each other, to discover new ways to be creative, and to realize that the performance was unique, in part, because of their contributions to it. To evaluate the project's success, we sought feedback from audiences at five performances at the Adrienne Arsht Center for the Performing Arts in Miami (December 2007) and at four performances at the 01SJ Festival in San Jose (June 2008). Audiences were invited to complete written surveys, which included Likert-style questions and space for free-response comments. In Miami, 50 audience members submitted surveys out of approximately 400 total during the run; in San Jose, 25 audience members submitted surveys out of approximately 300 total during the run. As Table 1 shows, audiences at both performances had limited exposure to experimental music and little formal musical training. In that regard, they were exactly the types of people we wished to attract to the performance and creatively engage during the show.

Audiences in Miami did have fun (interpolated median 4.61, where 5 indicates strongest agreement) and enjoy participating (4.58), but they were divided over whether they had been creative (3.40) and whether the performance would have been different without them (3.20). We did observe much creativity, including spontaneous ballroom dancing and conga lines, but these activities were usually initiated by a handful of audience members, with other participants following those leaders. We also noted that audience members were most creative when only a handful of participants were on stage; they were much less likely to initiate creative movement when dozens of people were participating.

Miami audiences also expressed some concern and confusion about their role; only some of them understood how their actions shaped the performance (3.58). In written comments, one person never realized that "I didn't have to do just what the dancers told me to – that I could move around." Another complained that it "felt very unorganized, chaotic, unstructured." Several respondents suggested the addition of a verbal introduction and explanation prior to the performance to help clarify the interaction and their role. These sentiments were reinforced during informal question-and-answer sessions we held after each performance.

Using this feedback, we made significant revisions to *Flock* before presenting the work in San Jose. We developed a new video animation (see section 3.4 above) that more directly rendered the position and notation data, to facilitate the audience's understanding of the work's underlying processes. We added a short introduction before the performance, in which we explained and demonstrated the basic mappings from position data into notation, video, and electronic sound. We encouraged the dancers to take a more active role as facilitators, making more spontaneous decisions about how to instigate creative audience behavior rather than handing out instruction cards (see section 3.5.2 above). And we further reduced the number of simultaneous participants on stage, creating a more intimate participatory environment in which audience members were more likely to participate creatively.

San Jose audiences responded to all but one survey question more positively than Miami audiences had. While revisions may not have been the only factor — our San Jose audiences were younger and more tech-savvy than in Miami — we do think the revisions played a significant role in this shift. Among the biggest improvements were in questions about their connection to the music (+0.86) and their belief that the performance would have been different without them (+0.72). They also felt significantly more creative (+0.40). They felt less connected to the rest of the audience (-0.07), but we believe this was due to the cramped performance space in San Jose and the awkward audience seating arrangement it necessitated.

There remains room for much further improvement. For example, San Jose audiences complained that there was a noticeable delay — about a second — between their movement and the corresponding change in the music and video; this can make it harder for audience members to follow their role in the music and can discourage faster movement through the space. This latency is not so much a technical limitation as a design problem: music notation needs to be displayed to musicians in advance of when they read it, and so the video and electronic sound must be similarly delayed so as to remain in sync. We continue to search for a design that will enable us to minimize this latency without making the musicians' sight-reading task impossible.

4.1.1. Musicians

While developing the real-time music notation for *Flock*, we consulted extensively with saxophonists, and we worked closely with the members of the Miami Saxophone Quartet who premiered the work, using their feedback to do everything from adding and removing visual elements to changing colors and font sizes. The resulting notation was fairly clear and readable, given the constraints of small-screen display and real-time rendering.

With practice, saxophonists have grown comfortable reading *Flock*'s unique notation, and their own creative voices have emerged in performance. Saxophonist Jason Kush, who performed *Flock* in Miami, noted that once he became more comfortable reading the notation, he began to develop a more diverse vocabulary of musical gestures to respond to it, taking it "less literally…[and]…painting more of a sonic picture of what was up there." The resulting music, he felt, became more cohesive as he learned to better read ahead in each measure to form phrases, and as he learned the vocabulary of the other musicians and could start to predict what they would play (personal communication, 15 February 2008).

The biggest challenge facing musicians in *Flock* is that there is simply too much to do. The saxophonists must sight-read an often dense and complex score. They must coordinate the music they play with the other musicians and sometimes with electronic sound. They must read textual cues, move around the performance space, and avoid collisions with other participants. And they must do all of this while focusing their eyes on a tiny display affixed to their instrument.

Most saxophonists who have performed *Flock* tell us that their focus on the music — both mentally and visually — precludes much consideration of movement. They slowly wander around, moving a few steps at a time, but their movements are not creative decisions designed to shape the music notation they see and the actions of other participants. They are movements planned so as to avoid colliding with anyone else.

If budget allows in future performances, we would like to experiment with eyepiece or eyeglass-based displays. Such displays could enable the musicians to view the music notation superimposed onto their normal field of vision; they would no longer have to focus on a screen at the expense of seeing the rest of the performance space.

4.1.2. Dancers

Dancers were added to *Flock* late in the development of the work, based on our experiences with informal test presentations of draft versions of the piece. Their incomplete integration into the work remains evident. At each performance of *Flock*, we have recruited enthusiastic and talented dance students, and through limited rehearsal time and a largely improvisational approach, they have contributed tremendously to the success of the work. Yet their role remains somewhat vague and haphazard; the dancers have a significant influence over the audience's participation, the music, and the pacing of the performance, but they have only slightly more familiarity with the work's design and structure than the audience. In future performances, we would like to collaborate with a choreographer or theatrical director who can help to give clarity and structure to the dancers' participation, guide a more extended rehearsal process, and contribute a unique perspective towards the larger structure of the performance and the dancers' role within it.

4.1.3. Future Work

Our technical systems for *Flock* are robust, but we continue to develop the interactive framework and structure as we prepare for upcoming performances, seeking a balance between organized and free movement, between more transparent mappings and more compelling music; and between the audience's role as participants and spectators. These are challenging problems, in particular because experimentation often necessitates the presence of a full audience. But this is also part of the reason this project excites us: each performance brings fresh people, new social dynamics, and new insights.

5. ACKNOWLEDGEMENTS

Flock was commissioned by the Adrienne Arsht Center for the Performing Arts in Miami, with additional support from the Funding Arts Network, the Georgia Tech Foundation, and Georgia

Tech's GVU Center. It was premiered in Miami in December 2007. Those performances, which were produced by iSAW, featured saxophonists Gary Keller, Gary Lindsay, Jason Kush, and Mike Brignola, and dancers Michael John Harper, Michael Fernandez, Lauren Wojcik, and Kamaria Dailey. Frank Dellaert, Martin Robinson, Dan Hou, and Justin Berger developed additional software. Liubo Borissov created the original video animation. Additional thanks to Steve Dietz and the Rova Saxophone Quartet for presenting the piece at 01SJ in June 2008.

Performance videos, a score, and the Flock Vision Toolkit are available at http://www.jasonfreeman.net/flock/.

6. REFERENCES

Assayag, G., Rueda, C., Laurson, M., Agon, C., and Delerue, O., 1999. Computer-Assisted Composition at IRCAM: From PatchWork to OpenMusic. *Computer Music Journal,* 23(3), 59-72.

Baird, K., 2005. Real-Time Generation of Music Notation via Audience Interaction Using Python and GNU Lilypond. *In: Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME),* 26-28 May 2005 Vancouver, 240-241.

Brown, E., 1961. *Available Forms I.* New York: Schirmer.

Carpenter, L. and Carpenter, R., 1999. Audience Participation. *In*: T. Druckrey, ed. *Ars Electronica: Facing the Future*. Cambridge: MIT Press, 395-396.

Chénard, M., 2000. The Millenium Symphony: A Work for the Beginning of Time, Part I: The Musical Challenge. *La Scena Musical,* 5(8), 12–15.

Cope, D., 1996. *Experiments in Musical Intelligence.* Middleton: A-R Editions.

Didkovsky, N., 2004. Recent Compositions and Performance Instruments Realized in the Java Music Specification Language. *In: Proceedings of the 2004 International Computer Music Conference*, 1-6 November 2004 Miami, 746-749.

Didkovsky, N. and Burk, P., 2001. Java Music Specification Language, an introduction and overview. *In: Proceedings of the 2001 International Computer Music Conference*, 17-22 September 2001 Havana, 123-126.

Do, E. and Gross M., 2007. Environments for Creativity – A Lab for Making Things. *In: Proceedings of the 6th ACM Conference on Creativity and Cognition*, 13-15 June 2007 Washington, D.C., 27-36.

Feldmeier, M., and Paradiso, J., 2007. An Interactive Music Environment for Large Groups with Giveaway Wireless Motion Sensors. *Computer Music Journal,* 31(1), 50-67.

Flintham, M., Benford, S., Anastasi, R., Hemmings, T., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., and Row-Farr, J., 2003. Where on-line meets on the streets: experiences with mobile mixed reality games. *In: Proceedings of the SIGCHI conference on Human factors in computing systems*, 5-10 April 2003 Fort Lauderdale, 569-576.

Freeman, J., Ramakrishnan, C., Varnik, K., Neuhaus, M., Burk, P., and Birchfield, D., 2005. Auracle: A Voice-Controlled Networked Sound Instrument. *Organised Sound,* 10(3), 221-231.

Freeman, J., 2008a. Glimmer: Creating New Connections. *In*: R. Adams, S. Gibson, and S. Arisona, eds. *Transdisciplinary Digital Art: Sound, Vision and the New Screen*. Berlin: Springer, 270-283.

Freeman, J., 2008b. Extreme Sight-Reading, Mediated Expression, and Audience Participation: Real-Time Music Notation in Live Performance. *Computer Music Journal,* 32(3), 25-41.

Garton, B., and Topper, D., 1997. RTcmix — Using CMIX in Real Time. *In: Proceedings of the 1997 International Computer Music Conference (ICMC)*, 25-30 September 1997 Thessaloniki.

Gutknecht, J., Clay, A., and Frey, T., 2005. GoingPublik: Using Realtime Global Score Synthesis. *In: Proceedings of 2005 International Conference on New Interfaces for Musical Expression (NIME)*, 26-28 May 2005 Vancouver, 148-151.

Hasse, J., 1986. *Moths*. Euclid: Visible Music.

Hiller, L. and Isaacson, L., 1959. *Experimental Music: composition with an electronic computer*. New York: McGraw Hill.

Jordà, S., Geiger, G., Alonso, M., and Kaltenbrunner, M., 2007. The reacTable: Exploring the Synergy Between Live Music Performance and Tabletop Tangible Interfaces. *In: Proceedings of the First International Conference on Tangible and Embedded Interaction (TEI07)*. 15-17 February 2007 Baton Rouge, 139-146.

Kim-Boyle, D., 2005. Musical Score Generation in Valses and Etudes. *In: Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME)*, 26-28 May 2005 Vancouver, 238-239.

Levin, G. and Lieberman, Z., 2004. In-Situ Speech Visualization in Real-Time Interactive Installation and Performance. *In: Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering*, 7-9 June 2004 Annecy, 7-14.

Maynes-Aminzade, D., Pausch, R., and Seitz, S., 2002. Techniques for Interactive Audience Participation. *In: Proceedings of the IEEE International Conference on Multimodal Interfaces*, 14-16 October 2002 Pittsburgh, 15-20.

McAllister, G., Alcorn, M., and Strain, P., 2004. Interactive Performance with Wireless PDAs. *In: Proceedings of the 2004 International Computer Music Conference (ICMC)*, 1-6 November 2004 Miami, 702-705.

Nummiaro, K., Koller-Meier, E., and Van Gool, L., 2003. An Adaptive Color-Based Particle Filter. *Image and Vision Computing,* 21(1), 99-110.

Pelletier, J.M., 2008. *CV.JIT: Computer Vision for Jitter* [online]. Available from: http://www.iamas.ac.jp/~jovan02/cv/ [Accessed 15 December 2008].

Sauter, J. and Lüsebrink, D., 2000. *BodyMover* [online]. Available from: http://www.artcom.de/images/stories/2_pro_bodymover/bodymover_e.pdf [Accessed 15 December 2008].

Schmit, T., 1962. *Sanitas No. 35* [online]. Available from: http://www.artnotart.com/fluxus/tschmit-sanitasno.35.html [Accessed 15 December 2008].

SpliceMusic, 2008. *SpliceMusic* [online]. Available from: http://www.splicemusic.com [Accessed 15 December 2008].

Snibbe, S., 1998. *Boundary Functions* [online]. Available from: http://www.snibbe.com/scott/bf/index.htm [Accessed 15 December 2008].

Stockhausen, K., 1957. *Klavierstück XI*. London: Universal Edition.

Wakefield, G., and Smith, W., 2007. Using Lua for Audiovisual Composition. *In: Proceedings of the 2007 International Computer Music Conference (ICMC)*, 27-31 August 2007 Cophenhagen.

Wulfson, H., Barrett, G., and Winter, M., 2007. Automatic Notation Generators. *In: Proceedings of the 2007 International Conference on New Interfaces in Musical Expression (NIME)*, 6-10 June 2007 New York, 346-351.

Xenakis, I., 1992. *Formalized Music*. Stuyvesant: Pendragon Press.

Zbyszynski, M., Wright, M., and Campion, E., 2007. Design and Implementation of CNMAT's Pedagogical Software. *In: Proceedings of the 2007 International Computer Music Conference (ICMC)*, 27-31 August 2007 Copenhagen, 57-60.

About the Authors

Jason Freeman's works break down conventional barriers between composers, performers, and listeners, using cutting-edge technology and unconventional notation to turn audiences and musicians into compositional collaborators. His music has been performed by the American Composers Orchestra, Speculum Musicae, the So Percussion Group, and the Rova Saxophone Quartet, among others. Freeman received his B.A. in music from Yale University and his M.A. and D.M.A. in composition from Columbia University. He is currently an assistant professor in the School of Music at Georgia Tech in Atlanta.

Mark T. Godfrey received an MS in music technology from Georgia Tech, where his research interests included music modeling and information retrieval, machine listening, and interactive human-machine musical performance. He also holds a BS and MS in electrical engineering and presently works for a startup developing music applications for mobile phones. He is an active musician and enjoys performing and recording music.