Orchestral Performance Companion: Using Real-Time Audio to Score Alignment

Matthew Prockup, David Grunberg, Alex Hrybyk, and Youngmoo E. Kim
Drexel University

The Orchestral Performance Companion system guides listeners through orchestral performances with an automated personal guide that presents time-relevant annotations and contextual information on a handheld device.

Modern-day museums often provide visitors with an automated, handheld personal tour guide, usually in the form of an audio recording that includes facts about individual exhibits. This helps museum visitors better understand and appreciate what they are viewing and lets exhibit curators communicate with audiences on a personal level. Similarly, some people who attend orchestral performances find it difficult to follow and understand a concert program during the performance and would benefit from additional guidance. However, in this case, spoken audio descriptions are not an option.

Many performing arts organizations are actively experimenting with interactive audience education activities. These activities can include lectures preceding each performance that introduce the program to audiences at a broadly accessible level. Projected visual information (photos, graphics, and text) may also be presented during the concert. However, this constrains the presentation to one broad stream of information for everyone and requires constant attention and control input from a human operator.

Our system, the Orchestral Performance Companion, provides an automated personal guide that presents time-relevant annotations and contextual information on a handheld device. Using this system, concert attendees can learn about the music as it is performed and gain a deeper understanding of each piece. In addition, because it is viewed on a personal level, it is unobtrusive to audience members who do not wish to use the system. The system automatically determines the current location within a live performance, in terms of the music measure number, without the need for a human score follower to trigger events. To accomplish this, the system compares the live audio with a previously analyzed recording of the same composition. The computed position is transmitted wirelessly to the audience's handheld devices, which then display information that is appropriate for the current position. The annotations for each piece are prepared ahead of time in collaboration with the performing ensemble and other musical experts.

Educating the Audience

Several ongoing efforts are seeking to improve the accessibility of classical music. For example, major orchestras have presented special performances, entitled *Beyond the Score* (originally developed by the Chicago Symphony Orchestra), to better inform audiences about classical music. This effort uses multimedia in a large performance setting to communicate with and educate audiences. These concerts begin as lectures, with the conductor leading the orchestra through live excerpts of the music and discussing its features with the audience, aided by actors and projected visuals. For example, the Beyond the Score performance of Bartok's *The Miraculous Mandarin* incorporates drawings of urban landscapes and gritty scenery displayed on a theater screen, while a narrator and an actor discuss the plot. When they describe the character of a poor girl, symbolized by clarinets, several images of such a girl are displayed and the clarinets perform the corresponding theme from the piece. The performance includes a narrator reading letters by Bartok and pictures of his manuscript. After the lecture, when the audience has a deeper understanding of what

1070-986X/13/$31.00 © 2013 IEEE

Published by the IEEE Computer Society

Figure 1. Overall system design of the Orchestral Performance Companion. The system uses acoustic features extracted from the live music stream and aligns them with those extracted from a previous recording of the same piece. A user's mobile device displays time-relevant contextual information related to the aligned position.

they are about to hear, the orchestra performs the piece in its entirety.

Events such as Beyond the Score address the entire audience as a whole, but such presentations may not appeal to all concertgoers. The introduction of powerful mobile devices (such as handheld computers and smartphones) makes a more personal interactive and educational experience possible. In 2004, the Kansas City Symphony first implemented such a concept, using a personal digital assistant (PDA) to convey contextual information during a live performance. Their Concert Companion software displayed pictures, text, and small video clips pertaining to the live music on an iPaq PDA, but it was triggered by a human operator following a music score to cue the updates on the devices.

System Design

The Orchestral Performance Companion focuses on live orchestral performances. Figure 1 outlines the overall system design. As a performance takes place, live audio is streamed to a computer that attempts to locate the orchestra's current position within a piece. The system uses acoustic features extracted from the live music stream, aligning them with those extracted from a previous recording of the same piece. This effectively determines the position in the live performance as it relates to a corresponding position in a reference. Because the temporal locations of measures and other events in the reference are already known, the system can determine the position of the live performance (for example, "measure 325"). The position is then sent to the handheld devices, which display information relevant to the current location within the piece.

Audio Alignment

The system effectively follows a performance by aligning the live music with a previously annotated reference recording. This alignment is performed using an acoustic feature known as a *chroma*, which is computed for both the live audio and reference recording. Using dynamic time warping (DTW) in conjunction with the chroma features, it is possible to determine a time path through the reference recording that most closely aligns with the incoming live audio.

Prior to a live performance, a reference recording for each programmed piece is manually time stamped (annotated with time values corresponding to important measure numbers). When the tracking application determines the equivalent position of the live audio as related to the reference recording, it looks up the corresponding measure number, which is then broadcast to the audience's handheld client devices to trigger the display of relevant content.

Chroma Features. A chroma represents the energy distribution of an audio signal according to the 12 pitch classes (A, A#, B, . . ., G, G#) of the Western music scale.[1] Each pitch class corresponds to a set of fundamental frequencies. The pitch classes' frequency centers (in hertz) are logarithmically spaced, repeating with every octave (a doubling of the fundamental frequency). Table 1 provides the fundamental frequencies of each pitch class across multiple octaves. Chroma can be estimated by analyzing

Table 1. Central frequencies of pitch classes over multiple octaves (in hertz).

Note	Octave 1	Octave 2	...	Octave 5	Octave 6
A	55.0	110.0		880.0	1,760.0
A#	58.2	116.5		932.3	1,864.7
B	61.7	123.5		987.8	1,975.5
C	65.4	130.8		1,046.5	2,093.0
C#	69.3	138.6		1,108.7	2,217.5
D	73.4	146.8	...	1,174.7	2,349.3
D#	77.8	155.6		1,244.5	2,489.0
E	82.4	164.8		1,318.5	2,637.0
F	87.3	174.6		1,396.9	2,793.8
F#	92.4	185.0		1,480.0	2,960.0
G	98.0	196.0		1,568.0	3,136.0
G#	103.8	207.7		1,661.2	3,322.4

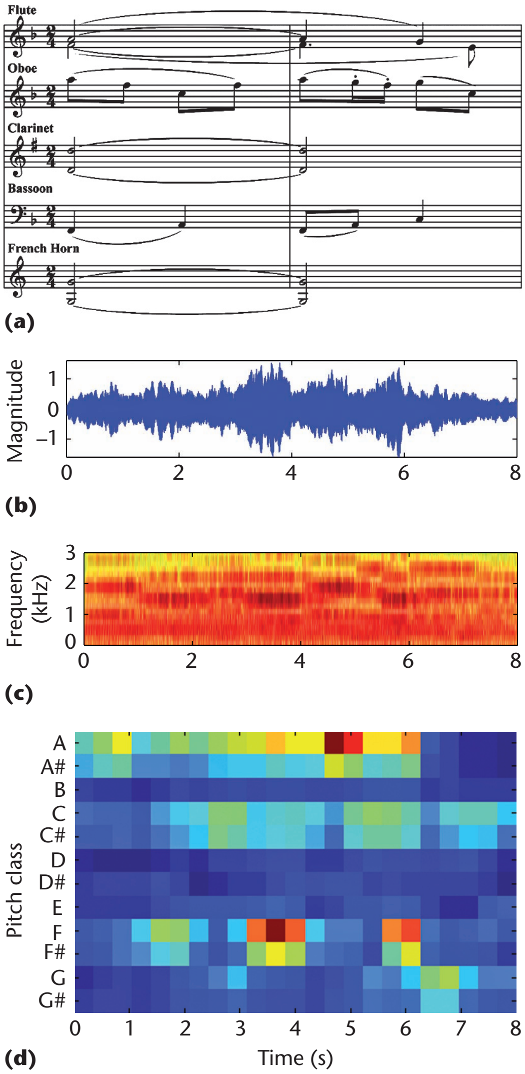

Figure 2. Chroma feature calculation. This example shows the (a) score, (b) audio waveform, (c) spectrogram, and (d) chromagram for two measures of Brahms's Violin Concerto, Movement 2.

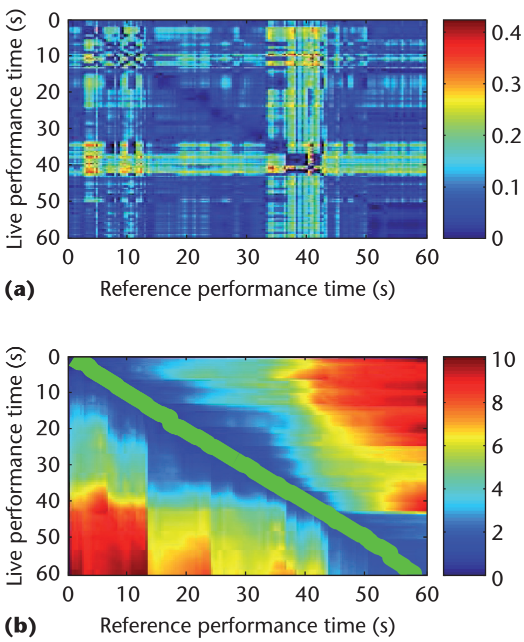

Figure 3. Dynamic time warping. (a) The distance and (b) cost matrices with the alignment path for two performances of Brahms' Violin Concerto.

the spectral content of a short time window (frame) of audio and filtering frequencies according to individual pitch classes. By combining across octaves (for example, summing the energies near 220, 440, or 880 Hz), we compute the total contribution of each pitch class (indicated by fundamental frequency) in each frame of the signal.

This is not a transcription of the music; the octave, or pitch height, of the notes is lost. Furthermore, all musical instruments produce upper partials (frequencies at integer multiples of the fundamental), but contributions of these partials may be attributed to a different pitch class from the played note. For example, the second partial of the fundamental frequency 220 Hz (chroma A) is 660 Hz, which will be counted toward chroma E. However, sections of audio containing similar notes will have similar chroma, so this feature is robust for aligning two performances of the same piece of music.[2,3] Figure 2 shows an example *chromagram*, which consists of frames of chroma over time.

Before a performance, we compute the chroma features of the reference recording at 1/3-second intervals. This rate is slow enough that the frames encompass full notes and fast enough to permit frequent position updates. Once the concert begins, the system computes chroma features from the incoming live music, also at 1/3-second intervals.

Time Alignment. Our system is based on a DTW algorithm. DTW is designed to align two signals with similar content but with an exhibiting time variation with respect to one another.[4] The algorithm is robust with respect to minor differences between performances, such as the addition or omission of some notes,[5] and has been used successfully in other music score-following[6] and score-alignment applications.[7] DTW has been used to align sets of chroma features to detect "covers" (different performances) of the same composition[2] and was a good starting point for our application.

In a performance setting, the chroma features from the live audio are computed and appended to the input chromagram.[8] The tracking application computes the difference between each reference chroma and live chroma frame (Figure 3a). It then identifies causal paths from the beginning of the

Table 2. System tracking performance using other performances and MIDI as the reference for alignment.

Artist	Piece	Type	Period	Recordings	Average length (min)	Accuracy (%)**
Bach	*Orchestral Suite No. 3*	Acoustic	Baroque	2	19.4	98.80
Brahms	*Violin Concerto*	Acoustic	Romantic	3	40.8	92.29
Handel	*Royal Fireworks Suite**	Acoustic	Baroque	2	6.1	95.64
Haydn	*Symphony No. 94*	Acoustic	Classical	2	22.2	92.74
Mozart	*Clarinet Concerto*	Acoustic	Classical	2	28.4	95.86
Mozart	*Violin Concerto No. 5*	Acoustic	Classical	3	26.7	96.16
Shostakovich	*Symphony No. 11*	Acoustic	Modern	3	59.7	81.61
Strauss	*Don Juan*	Acoustic	Romantic	3	16.7	91.20
Strauss	*Don Quixote*	Acoustic	Romantic	3	47.5	94.62
Live total				23	1275.6	93.21
Bach	*Orchestral Suite No. 3*	MIDI	Baroque	2	18.0	98.30
Handel	*Royal Fireworks Suite**	MIDI	Baroque	2	5.1	99.83
Haydn	*Symphony No. 94*	MIDI	Classical	2	22.6	94.05
Mozart	*Clarinet Concerto*	MIDI	Classical	2	28.1	97.61
Mozart	*Violin Concerto No. 5**	MIDI	Classical	3	17.1	93.77
Strauss	*Don Juan*	MIDI	Romantic	3	16.3	76.08
MIDI total				14	248.7	93.27
Total				37	1560.3	93.23

* Excerpts

** Accuracy is defined as being within two measures of the true position.

difference matrix to each cell in the matrix, sums the differences along each path, and compares the sums to find the minimum cumulative difference (or cost) to reach any cell (Figure 3b). A low cost for a cell implies that the live and reference chroma along the best path are similar, and thus that path closely aligns the two performances.

The process of calculating the distances and then the cost between the two sets of features is a key feature of DTW.[9] Although the algorithm was traditionally run offline, recent implementations have shown that its principles can be effectively used in online systems as well.[10] Thus, we update the cost matrix after each new frame of live audio is detected. We locate the minimum-cost point at the current time index in the live performance and use that point to identify the closest frame in the reference recording to the current frame of live audio.

Our system also differs from traditional DTW in that, once the current position is found, no back tracing is necessary to fully align the live and reference performances. Our application only requires the current position, so a full alignment is superfluous. By eliminating unlikely paths, we can improve the system's efficiency to the point where it can run in real time on a modern CPU such as an Intel Core 2 Duo.[10] For instance, we might not consider alignments that would result if the tempo of the live performance were more than twice as fast as the reference.[11]

Tracking System Evaluation

We have evaluated the performance of our tracking algorithm on more than 20 hours of recordings of orchestral works spanning several musical periods. Each of the pieces used for the tracking system evaluation were part of the beta test broadcasts. We aligned every performance of each piece against all other versions of that piece and determined the percentage of each piece identified as being within two measures of its actual position by the tracking system. The system is accurate to within two measures of the true position more than 90 percent of the time. Table 2 provides detailed performance results.

Figure 3b shows an example time alignment. This plot depicts the time alignment for an excerpt of two performances of Brahms'

Violin Concerto. The slope of the alignment represents the ratio of the reference and live performance tempos. When this ratio is constant over an excerpt, the line should be highly linear (as in Figure 3b). When the ratio changes, the line should change slope. A proper alignment will therefore be a series of piecewise linear segments, each segment corresponding to the ratio of the two tempos. If the alignment ceases, the line will become flat. This may occur during improvised sections in the work called *cadenzas*, which can vary greatly from performance to performance, preventing alignment.

Our system also performs well when using audio rendered from musical instrument digital interface (MIDI) files as the reference for alignment. MIDI is a symbolic representation (similar to a score) used to trigger synthetic instruments. MIDI files include precise measure times in their representation, eliminating the need to laboriously label the measure times in a reference recording by hand. The resulting performance when tracking MIDI files against five of the pieces in our dataset again exceeds 90 percent accuracy overall. Table 2 gives the full results.

System Implementation

In 2009 and 2010, we partnered with Specticast to present theater-style live broadcasts of Philadelphia Orchestra concerts in a 300-seat auditorium on the campus of Drexel University. The Specticast service is similar to the popular "Metropolitan Opera: Live in HD" broadcasts that are shown in movie theaters around the world. The efforts diverge, however, in terms of both target audiences and core technology. Specticast is primarily marketed toward groups that may have difficulty attending performances, such as those in assisted-living communities and community centers, and it emphasizes portability and ease of use.

For this reason, Specticast uses Internet streaming for their events. (The Metropolitan Opera uses a dedicated digital satellite channel for their presentations.) Furthermore, Specticast presentations require only a single receiver box and a standard broadband Internet connection as an input, which outputs a high-definition multimedia interface (HDMI) video/audio signal. A data stream of approximately 6 megabits per second (Mbps) is sufficient for a full HD (720 pixel) picture suitable for an auditorium or small theater.

The Orchestral Performance Companion system was developed and refined over the five Specticast concert programs available during the 2009–2010 season. These five concerts acted as beta tests for our system. The Philadelphia Orchestra has since adopted the system, allowing its use live in the concert hall for testing during the 2010–2011 and 2011–2012 seasons and in full public deployment in 2012–2013.

Personal Client Device

The goal of the Orchestral Performance Companion is to enhance the classical concert experience for a range of audiences, so the client device and application must exhibit the following attributes:

- The application must be highly intuitive and easy to use, presenting annotations that are clear and easily viewed.
- The devices should be unobtrusive to other audience members.
- Communication with the tracking server should be transparent so that the client receives tracking updates automatically.

Our system was developed using the iOS platform (for Apple iPhone and iPod Touch devices). The capabilities and popularity of these mobile devices provide an ideal platform for the Orchestral Performance Companion.

Our primary interface uses a slideshow-style display (see Figure 4a). As the music progresses, pages of information containing text and supporting images are displayed. Users can allow the system to update pages on its own, guiding them through a performance, or they can page through the information at their leisure. Once a user chooses to page manually, a banner appears notifying them that they are no longer viewing the currently relevant information. They can return to the live position by tapping the banner. In addition, the current position within the piece is visible at all times via an updating timeline at the bottom of the slide. Tick marks within the timeline show the positions of annotations.

The application is designed to easily accommodate supplemental information such as musical vocabulary and definitions. In Figure 4a, the word "fluttertonguing" is a hyperlink to the glossary, which offers a definition (Figure 4b). In this way, users can tap any highlighted term for a definition and then easily return to their previous position in the music when they are finished.

Annotations can also be presented as multiple tracks of information, each focusing on a different aspect of the music. These different tracks can be presented on slides, as before, or using an alternate interface. This second interface shows a "roadmap" for the music. The beginning and end of each piece are depicted as starting and ending points on a map. Users can glance down at our application's map view to obtain a sense of the current location within the overall structure of the piece—where it's been and where it's heading. The example map in Figure 4c shows the entrance of a new section. Users can easily switch between the map and slide information views via buttons at the bottom of the interface screen.

We believe that offering such supplemental information interactively will help guide audiences through a performance by allowing them to make choices in terms of content, layout, and depth of presentation.

Operator Control

The operator control panel displays real-time data so that an operator can easily survey system performance and status. The panel contains two live updating figures that show a detailed view of the least-cost path through a short time window surrounding the current music position as well as the overall path from the beginning of the performance to the current position. The system can be configured so that it begins tracking from any measure in the piece; this allows for real-time adjustments in response to unpredictable events (such as solo cadenzas or the repetition or omission of entire sections, when such repetition is optional). A full concert program is loaded prior to a performance, allowing for easy transition between movements and pieces.

Tracking Server

The server has two primary functions: it provides the devices with the annotation data for the concert before it begins and it broadcasts the live position in the piece to the handheld clients as the piece is being performed.

Figure 4. Annotated content relating to measure 220 of Don Quixote *by Richard Strauss. (a) The yellow words on the slide-style interface (such as "dissonant," "tremolo," and "fluttertonguing") are hyperlinks to a glossary of musical terms. (b) Users can look up the highlighted words in the application's glossary. (c) The map-style interface shows the beginning and end of each piece.*

When audience members first launch the application on their client device, the annotated content for the concert is downloaded in full so that the server is not overloaded during the performance with requests for large amounts of image and text data simultaneously.

Once the annotation data is loaded and the concert begins, the clients receive performance position updates (in terms of measure numbers) from the tracking server. In our initial implementation, the server sent data packets to individually connected clients, but this was not scalable. The most recent system uses User Datagram Protocol (UDP) packets broadcast over a multicast address. This lets all the client

Table 3. Collaborating musicologists who developed annotations for programs presented in 2009–2012.

Musicologist	Annotated program
Myron Moss Drexel University	*Clarinet Concerto* Wolfgang Amadeus Mozart
Daniel Fox University of Pennsylvania	*Symphony No. 11* Dmitri Shostakovich
Leonardo Dugan The Philadelphia Orchestra	*Don Juan* Richard Strauss *Don Quixote* Richard Strauss *Violin Concerto No. 5* Wolfgang Amadeus Mozart
Makiko Freeman and staff The Philadelphia Orchestra	*Capriccio Espagnol* Nikolai Rimsky-Korsakov *Liebeslied for Violin and Orchestra* Friedrich Kreisler *Liebesfreud for Violin and Orchestra* Friedrich Kreisler *Meditation from Thais, for Violin and Orchestra* Jules Massenet *Toccare from Violin Concerto* John Adams *Overture to William Tell* Gioachino Rossini *A German Requiem* Johannes Brahms *Rhapsody on a Theme of Paganini* Sergei Rachmaninoff *Symphony No. 2—"Little Russian"* Pyotr Tchaikovsky
Jennifer Higdon Composer, Curtis Institute of Music	*Concerto for Orchestra* Jennifer Higdon

devices listen to a common multicast IP address and all receive position update information simultaneously. The router keeps track of multicast group members. When the group IP receives data, the router forwards it to all other multicast group members. This is viewed as a connectionless protocol in which the server need not keep track of all the clients, which can be cumbersome with a great number of users. The only limitation on the number of clients is the number supported by the hardware (router and access points) supplying the wireless internet connection.

Content Authoring

In addition to its technical design, one of the most important system components is the contextually relevant content provided by the application to concert audiences. These presentations are developed and curated prior to the performance. The content itself is placed into a server database linking the multiple streams of annotations to measure numbers. The tracking component relays the current position in terms of measure number so the client application can display time-appropriate content to the user.

Before each concert, the musicologists begin creating time-relevant performance notes linked to measure numbers or rehearsal markings in the music. Topics range from music theory concepts to a piece's historical significance, spanning anything the author feels may be helpful for an audience member to better understand a performance.

These collaborations generate greater value than just the content created. In working closely with the annotators, we have obtained feedback regarding our system from those who are most likely to incorporate it into their own educational activities. We have integrated this feedback into our system to improve the presentation of the content that the designers seek to convey.

From inception through full integration, the content for each program takes about two weeks of part-time work to prepare. We have worked with multiple partners to develop annotation content, including members of the Philadelphia Orchestra staff and one of the composers. Table 3 lists our collaborations with musicologists.

Philadelphia Orchestra Integration and System Testing

Drexel University and the Philadelphia Orchestra have teamed up to present these enhanced performances during subscription concerts in the 2011–2012 and 2012–2013 seasons. We have worked closely with them in performing live concert beta tests with test audiences as well as closed tests to investigate issues such as network scalability. In addition to the Drexel-run broadcast beta tests, four additional trial concerts were performed in the concert hall with a live orchestra.

In the fall of 2011, we performed a test in which iPod Touches were given to 20 Philadelphia Orchestra staff members during the performances of *A German Requiem* by Johannes Brahms and *Don Juan* by Richard Strauss.

Each member filled out an open-ended survey providing feedback about the interface, annotated content, possible cultural/environmental issues, and technical glitches. Users consistently praised the system's ease of use, the variety and usefulness of the content presented, and the positive and engaging experience the system offered. In addition, each member rated the system as a whole as poor, fair, good, or excellent. Out of the total ratings, 10 percent of users rated the system as ''excellent,'' 80 percent of users labeled the system as ''good,'' and 10 percent of users labeled the system as ''fair.'' There were no poor ratings.

In a live setting, a full hall of concertgoers attempting to access the internal Wi-Fi network can present issues if the network is not robust enough to handle a large number of clients simultaneously. It is also important that everyone sitting in the hall, regardless of location, has adequate network coverage. To support the network-load issues this project presents, we have worked closely with the Philadelphia Orchestra's Technology Infrastructure Department to reconfigure its internal network to allow for more Wi-Fi connections. In additional, we have increased the number of wireless access points, giving the concert hall more Wi-Fi coverage with greater signal strength.

After the installation, measurements were taken to confirm that every section in the concert hall had adequate signal strength. Another test involved a large number of clients sending and receiving wirelessly via a multicast group. This helped to confirm that the routers and wireless access points were configured properly to support a large-scale multicast session. The system was scaled to accommodate approximately 2,000 simultaneous users, which is the upper limit of the target audience.

Conclusion

We have developed a system for enhancing live performances to better inform and engage classical music audiences. The Orchestral Performance Companion is easy to use and takes advantage of the popularity and broad availability of iOS devices. The system, in its most recent state, was deployed at both live and broadcasted orchestral performance for audiences of various sizes.

On the technical side, we are working to improve network scalability and are experimenting with better ways of handling clients, allowing our broadcast server to communicate even more efficiently with a larger number of clients. In addition, we are striving to improve the system's usability as a whole. This includes streamlining the development of annotation content (which is still a tedious task), optimizing the application's design, and improving performance-tracking accuracy. As we continue to work with the Philadelphia Orchestra, these improvements will make the application easier for the orchestra to use and more informative and useful for large audiences. **MM**

Acknowledgments

We thank the staff members of the Philadelphia Orchestra for their continued collaborative efforts throughout this project. We also thank Specticast for its support for the concert broadcast events at Drexel. Additionally, we thank the John S. and James L. Knight Foundation. The project is supported through a Knight Arts Challenge grant, which has allowed Drexel University and the Philadelphia Orchestra to implement the project on a large scale.

References

1. M.A. Bartsch and G.H. Wakefield, ''To Catch a Chorus: Using Chroma-Based Representations for Audio Thumbnailing,'' *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,* IEEE CS, 2001, pp. 15–18.
2. D.P.W. Ellis and G.E. Poliner, ''Identifying 'Cover Songs' with Chroma Features and Dynamic Programming Beat Tracking,'' *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* vol. 7, IEEE CS, 2007, pp. IV-1429–1432.
3. J.H. Jensen et al., ''A Tempo-Intensive Distance Measure for Cover Song Identification Based on Chroma Features,'' *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* IEEE CS, 2009, pp. 2209–2212.
4. D.J. Berndt and J. Clifford, *Using Dynamic Time Warping to Find Patterns in Time Series,* tech. report WS-94-03, Assoc. Advancement of Artificial Intelligence, 1994.
5. R.B. Dannenberg, ''An On-line Algorithm for Real-Time Accompaniment,'' *Proc. Int'l Computer Music Conf.,* Computer Music Assoc., 1984, pp. 193–198.
6. R.B. Dannenberg and N. Hu, ''Polyphonic Audio Matching for Score Following and Intelligent Audio Editors,'' *Proc. Int'l Computer Music Conf.,* Computer Music Assoc., 2003, pp, 27–34.

7. J. Devaney and D.P.W. Ellis, ''Handling Asynchrony in Audio-Score Alignment,'' *Proc. Int'l Computer Music Conf.,* Computer Music Association, 2009, pp, 29–32.
8. S. Dixon, ''Live Tracking of Musical Performances Using On-line Time Warping,'' *Proc. 8th Int'l Conf. Digital Audio Effects,* 2005, pp, 92–97.
9. N. Orio and D. Schwarz, ''Alignment of Monophonic and Polyphonic Music to a Score,'' *Proc. Int'l Computer Music Conf.,* Computer Music Assoc., 2001, pp. 129–132.
10. R. Macrae and S. Dixon, "Accurate Real-Time Windowed Time Warping," *Proc. Int'l Soc. for Music Information Retrieval Conf.,* 2010, pp, 423–428.
11. E.J. Keogh and M.J. Pazzani, ''Derivative Dynamic Time Warping,'' *Proc. 1st SIAM Int'l Conf. Data Mining,* SIAM, 2001, pp, 150–159.

Matthew Prockup is doctoral student in the Music and Entertainment Technology Laboratory (MET-lab) at Drexel University. His research interests include human interaction and music. Prockup has an MS in electrical engineering and a minor in music theory and composition from Drexel University. He is a student member of IEEE. Contact him at mprockup@drexel.edu.

David Grunberg is a second-year graduate student in the Music and Entertainment Technology Laboratory (MET-lab) at Drexel University and is supported with a National Science Foundation Graduate Research Fellowship. His research interests include analysis of musical signals and applications for human-music interaction. Grunberg has an MS in electrical engineering from Drexel University. He is a student member of IEEE. Contact him at dgrunberg@drexel.edu.

Alex Hrybyk is a DSP research engineer at Lockheed Martin, specializing in streaming media and cybersecurity. Until 2010, he was an active member of the Music and Entertainment Technology Laboratory (MET-lab) at Drexel University. Hrybyk has an MS in electrical engineering from Drexel University. He is a student member of IEEE. Contact him at ahrybyk@drexel.edu.

Youngmoo E. Kim is an associate professor of electrical and computer engineering and the assistant dean of media technologies for the College of Engineering at Drexel University. His research group studies the machine understanding of audio. Kim has a PhD from the Massachusetts Institute of Technology Media Lab. He is a member of IEEE. Contact him at ykim@drexel.edu.