PHENICX: INNOVATING THE CLASSICAL MUSIC EXPERIENCE

Cynthia C. S. Liem[1] *Emilia Gómez*[2] *Markus Schedl*[3]

[1] Multimedia Computing Group, Delft University of Technology, c.c.s.liem@tudelft.nl
[2] Music Technology Group, Universitat Pompeu Fabra, emilia.gomez@upf.edu
[3] Department of Computational Perception, Johannes Kepler University, markus.schedl@jku.at

ABSTRACT

PHENICX ("Performances as Highly Enriched aNd Interactive Concert eXperiences") is an EU FP7 project that lasts from February 2013 to January 2016. It focuses on creating novel digital concert experiences, improving the accessibility of classical music concert performances by enhancing and enriching them in novel multimodal ways. This requires a user-centered approach throughout the project. After introducing the project, we discuss its goals, the technological challenges it offers, and current scientific and technological outcomes. Subsequently, we discuss how integrated prototypes combine several technological advances in the project into coherent user-ready interfaces, offering novel ways to experience the timeline of a concert, and rediscover and re-experience it afterwards. Finally, we discuss how PHENICX outcomes have been demonstrated live in concert halls.

***Index Terms*—** music information retrieval, video analysis, audio processing, information visualization, personalization, recommender systems, human-computer interaction, creative industries, classical music, concert

1. INTRODUCTION

Musical concerts typically are closed events: commonly, musicians perform in an enclosed physical space, playing for a privileged audience that acquired the right to be in that same enclosed space through purchased concert tickets. On one hand, under such a model, a social community will be physically formed within the concert venue. On the other hand, the walls of the concert venue put up a physical barrier, and at the same time, the local etiquette of the social community that identifies most strongly with the performed music puts up a social barrier. Because of these, people who would be interested in exploring live performances of unfamiliar music will be faced with an isolated, imposed and standardized concert situation they do not naturally identify with, thus remaining 'outsiders' to the music and its entourage.

This particularly becomes an issue in the case of Western classical music. While being a very strong example of European cultural heritage, with a history and performance tradition spanning several centuries in history, the genre suffers from very strong stereotypical audience images (an elitist and difficult music genre meant for old people). While the music is actually unconsciously being consumed at home and through the media by people who do not belong to this category (e.g. through film soundtracks and commercial music), these stereotypes prevent many of such people to actively go and attend a live classical music performance.

Therefore, the PHENICX project[1] investigates how technology can be used to create novel digital classical concert experiences, capable of engaging broader audiences for symphonic classical music. A broad research spectrum is addressed, investigating varying data sources and modalities, including audio, video, gesture and social data, and ranging from signal-based lower- and higher-level descriptors of musical pieces and their performances to user-oriented topics, such as personalization, recommendation and visualization techniques [1, **?**]. The academic topics are connected to several use cases [2], targeting the experience of a concert before, during and after its performance.

2. GOALS

The goal of the PHENICX project is to exploit state-of-the-art multimedia and internet technologies such that a live concert becomes a digital artefact. The artefact is *multimodal*: the concert experience is not only auditory. We combine input from other modalities, e.g. video, sensor data and text. Next to this, it is *multi-perspective*: we provide concert experiences from different viewpoints, both physical (location in the stage) and personal (different user profiles). Finally, it is *multilayer*: multiple layers of information are important to the musical experience, e.g. information about the musical piece, the composer, the performers, the instruments, or another performances of the same piece. As a result, novel engaging and interactive ways to explore live classical concerts are developed, which can enrich the concert experience before, during and after the concert. In working towards these goals, the project has the ambition to achieve a paradigm shift regarding the audience experience of concert performances,

[1] http://phenicx.upf.edu

978-1-4799-7082-7/15/$31.00 ©2015 IEEE

moving away from concert performances as isolated, exclusive and constrained events.

3. SCIENTIFIC AND TECHNOLOGICAL CHALLENGES

In the context of PHENICX, we have to face several scientific and technological challenges. In general, automated music processing technologies historically were strongly biased towards audio analysis, and towards popular music genres. In contrast, PHENICX focuses on Western classical music, which poses several data challenges that are not encountered for the majority of popular music.

For example, classical music pieces typically do not have a continuous tempo, a strongly present uniform beat, or clear audio onsets throughout the course of the piece. Instead, in symphonic music, many players (sometimes more than a hundred at once) play together at the same time, creating an intricate sound mixture. The structure of classical music pieces typically takes place at a larger scale, and with less obvious boundaries, than typical chorus-verse alternations as found in popular music. This means that audio analysis techniques based on popular music have to be adapted.

Furthermore, classical music pieces are usually more complex than popular music. Frequently, a piece consists of multiple movements, which are longer than the average popular song. The piece is attributed to a composer, who wrote the piece down in a score, and subsequently is performed and reinterpreted by musicians. This emphasizes the need to distinguish between pieces and their performances, and to align different performances to each other or to a score.

In a live concert, not only the hearing sense is addressed. During a performance, the audience can see musicians playing, and the very experience of being in a concert hall even addresses haptic and olfactory senses. This supports the view of a concert as a multimodal artefact, and thus for multimodal concert data processing techniques. Next to this, the social experience of a concert also is important: people may be more inclined to go to concerts they can attend or discuss with peers. On the other hand, we found that audiences of classical concerts are frequently rather reluctant to use digital online social platforms [**?**], which poses another challenge.

4. CURRENT OUTCOMES

In the following subsections, we discuss several research directions investigated in the project and point to recent achievements.

4.1. Multimodal data gathering

Treating a concert as a multimodal artefact requires offering sophisticated ways to manage and display synchronized hybrid data streams. RepoVizz[2] is a multimodal online database and visualization tool that can make these data streams more insightful and showcase different information visualizations [3]. In the PHENICX project, it integrates audio and video streams, digital score information, gesture data, manual annotations and computed descriptors.

4.2. Audio-to-Score alignment

Automatic alignment of classical performances to scores is challenging: structural differences may exist between different interpretations, and a performer will usually not maintain a constant tempo throughout a piece due to expressive and artistic choices. Both problems have been addressed in the project; new alignment methods were proposed to allow for effective alignment in the case of potentially omitted repeats [4] and technologies for live performance tracking and score-following were demonstrated to work very robustly on classical piano music, and now are being adapted to work with orchestra music [5]. Furthermore, work has been performed on performing note-level audio-to-score alignment, which is a useful resolution to consider when wishing to perform score-informed audio source separation [6].

4.3. Source separation

Being able to separate or emphasize voices in an orchestral audio mix is interesting for use cases in which a user would wish to focus on a particular part of an orchestra. In PHENICX, we have evaluated state-of-the-art approaches for sound source separation and, based on them, we are currently refining a novel method to emphasize different instrument sections of an orchestra, based on an informed source separation algorithm (requiring note-level alignment [6]) that can work on either mono or multi-channel recordings [3].

4.4. Audio and score description

The automatic description of audio and score information from a music piece facilitates ways to navigate and visualize relevant aspects of musical pieces and performances. Predominant melody is relevant as it is melody what humans imitate when singing along with music. Symphonic music presents a challenge to melody extraction methods [**?**] dealing with audio, as we analyze complex recordings where the melody is alternated between different instruments or instrument sections. In PHENICX, we have studied how humans agree when annotating the melody of symphonic music excerpts, and we have improved state-of-the-art methods for this particular repertoire [**?**]. Having an aligned scoreis useful to combine audio and score-based automatic description methods. In PHENICX, we do so to characterize the structure of

[2] `http://repovizz.upf.edu`

[3] listen to current separation results at `http://repovizz.upf.edu/phenicx/RCO-Eroica/`

musical pieces in terms of tonality [?], tempo, loudness or instrumentation (which e.g. can be used in *orchestra layout* visualizations[4]).

4.5. Visual analysis

As mentioned before, in an orchestra, multiple players will be active at the same time. While from audio, it is virtually infeasible to analyze instrumentalist playing behavior at the individual musician level, in the visual domain the different musicians are physically spaced apart, and as such can be segmented and analyzed. This becomes especially helpful when considering that professional concert videos typically are shot with multiple cameras, offering multiple parallel video data streams at the same time, and thus needing intelligent mechanisms to organize them. Within the PHENICX project, work is performed on automatically characterizing whether individual musicians are playing or not based on visual features. This information can then be used to improve several music information retrieval techniques[5]. As a first step, we used the results of our playing/non-playing detector for score-to-performance video alignment, even in case labels would be noisy or missing [7].

4.6. Gesture modeling

In an orchestra, the conductor plays a key role in the performance. In the scope of the PHENICX project, Sarasúa et al. [?][?] are working on exploring the possible applications that information captured from unobtrusive motion capture devices during musical performance can bring. These applications range from music description and visualization to real-time music interaction, using these devices as input for musical performance, as illustrated in Figure 1.

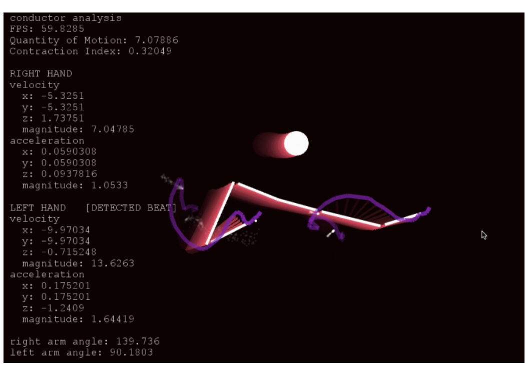

Fig. 1. Screen capture of the PHENICX demonstrator for visualizing conductor gestures.

[4]`http://repovizz.upf.edu/phenicx/RCO-Eroica/`
[5]`https://www.youtube.com/watch?v=6Oj70tvqo9c`

4.7. Social data mining

To reach new audiences and personalize the concert experience, we looked into data available through social media. A first insight we gained is that most fans of classical music are not very active on respective platforms [?]. Therefore, instead of performing extensive data crawling and machine learning experiments to create user profiles, we conducted a series of quantitative user studies to investigate correlates between aspects including demographics, musical education and knowledge, personality traits, genre preferences, and perception of musical segments according to emotions, tempo, complexity, and instruments. Initial results show that certain personality traits correlate to preferences in how listeners want to access or search for music [?], e.g. openness positively correlates with access by mood or emotion. Furthermore, we identified several correlations between personality traits and musical background [?], e.g. a positive correlation between extraversion and playing an instrument as well as attending concerts; a negative correlation between agreeableness and listening to non-classical music. Eventually, we will use these findings to personalize the information and visualizations provided to the concert audience.

5. USER-CENTERED APPROACH

To assess and verify the effectiveness and realism of proposed technological advances, active feedback of their envisioned audiences is essential. Therefore, from the start of the project, focus groups with user groups representing different consumer segments (experts, casual listeners and outsiders) have been held in The Netherlands, Spain and Austria. These focus groups offered insight in the initial perception of different planned use cases and corresponding technologies, early feedback on mock-ups of an integrated user experience, and feedback on proposed technological functionality.

6. INTEGRATED PROTOTYPE

Over the course of the project, an integrated tablet prototype is developed offering an enriched concert experience before, during and after a concert performance. The prototype, of which a screen capture is shown in Figure 2, currently offers a post-live experience integrating timed program notes, timeline commenting possibilities, score following, visualizations on orchestra activity, visual musician recognition, and the possibility to listen to separated instrumental sections.

7. LIVE DEMONSTRATIONS OF PHENICX

PHENICX outcomes have already successfully been demonstrated in concert halls. The Royal Concertgebouw Orchestra, one of the partners in the consortium, has supported tests of live functionality of the integrated prototype during a concert

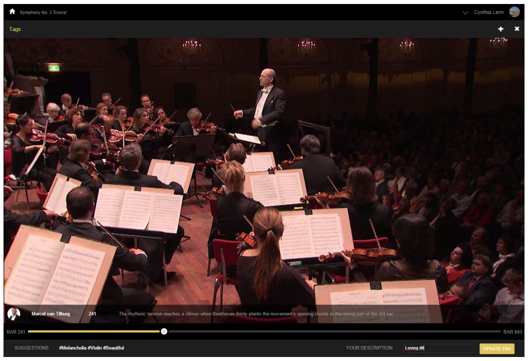

Fig. 2. Screen capture of the PHENICX integrated prototype.

Fig. 3. Test of live functionality of the PHENICX integrated prototype in the Amsterdam Concertgebouw.

(Figure 3). Furthermore, at the Singularity University Summit Spain, visualizations based on live analysis of performance data were displayed to the audience during an orchestra concert (Figure 4).[6]

8. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7 / 2007 - 2013 through the PHENICX project under Grant Agreement no. 601166.

9. REFERENCES

[1] E. Gómez et al., "PHENICX: Performances as Highly Enriched aNd Interactive Concert Experiences," in *Proc. SMAC/SMC*, Stockholm, Sweden, August 2013.

[2] M. Tkalčič, B. Ferwerda, D. Hauger, and M. Schedl, "Personality Correlates for Digital Concert Program Notes," in *Proc. UMAP*, Dublin, Ireland, June–July 2015.

[3] C.C.S. Liem et al., "Innovating the Classical Music Experience in the PHENICX Project: Use Cases and Initial User Feedback," in *Proc. WSICC*, Como, Italy, June 2013.

[4] M. Schedl and M. Tkalčič, "Genre-based Analysis of Social Media Data on Music Listening Behavior," in *Proc. ISMM*, Orlando, FL, USA, November 2014.

[6] http://phenicx.upf.edu/SUprometheus

Fig. 4. Live music visualizations during an orchestra performance at the Singularity University Summit in Spain, where we highlight which instruments are sounding in each moment.

[5] O. Mayor, "Web-based visualizations and acoustic rendering for multimodal data from orchestra performances using repovizz," in *Proc. WAC*, January 2015.

[6] Maarten Grachten, Martin Gasser, Andreas Arzt, and Gerhard Widmer, "Automatic alignment of music performances with structural differences," in *Proc. ISMIR*, November 2013.

[7] A. Arzt, S. Böck, S. Flossmann, H. Frostel, M. Gasser, C.C.S. Liem, and G. Widmer, "The Piano Music Companion," in *ECAI PAIS (System demonstration)*, August 2014.

[8] M. Miron, J.J. Carabias-Orti, and J. Janer, "Audio-to-score alignment at the note level for orchestral recordings," in *Proc. ISMIR*, October 2014.

[9] J. Salamon, E. Gómez, D.P.W. Ellis, and G. Richard, "Melody extraction from polyphonic music signals: Approaches, applications, and challenges," *Signal Processing Magazine, IEEE*, vol. 31, no. 2, pp. 118–134, 2014.

[10] J.J. Bosch and E. Gómez, "Melody extraction in symphonic classical music: a comparative study of mutual agreement between humans and algorithms," in *Proc. CIM*, Berlin, December 2014.

[11] A. Martorell and E. Gómez, "Contextual set-class analysis," in *Computational Music Analysis*, D. Meredith, Ed. Springer LNCS, In Press.

[12] A. Bazzica, C.C.S. Liem, and A. Hanjalic, "Exploiting Instrument-wise Playing/Non-Playing Labels for Score Synchronization of Symphonic Music," in *Proc. ISMIR*, October 2014.

[13] Á. Sarasúa and E. Guaus, "Beat tracking from conducting gestural data: A multi-subject study," in *Proc. MOCO*, Paris, France, 2014, ACM, ACM.

[14] Á. Sarasúa and E. Guaus, "Dynamics in music conducting: A computational comparative study among subjects," in *Proc. NIME*, London, United Kingdom, 06/2014 2014, Goldsmiths, University of London, Goldsmiths, University of London.

[15] B. Ferwerda, E. Yang, M. Schedl, and M. Tkalčič, "Personality Traits Predict Music Taxonomy Preferences," in *ACM CHI '15 Extended Abstracts on Human Factors in Computing Systems*, Seoul, Republic of Korea, April 2015.