

# Developing Music Harmony Awareness in Young Students through an Augmented Reality Approach

Federico Avanzini<sup>1</sup><sup>a</sup>, Adriano Baratè<sup>1</sup><sup>b</sup>, Mauro Cottini<sup>1</sup>, Luca Andrea Ludovico<sup>1</sup><sup>c</sup>  
and Marcella Mandanici<sup>2</sup><sup>d</sup>

<sup>1</sup>Laboratorio di Informatica Musicale, Dipartimento di Informatica “Giovanni Degli Antoni”,  
Università degli Studi di Milano, Via G. Celoria 18, Milano, Italy

<sup>2</sup>Dipartimento di Didattica della Musica, Conservatorio di Musica “Luca Marenzio”,  
Piazza A. Benedetti Michelangeli 1, Brescia, Italy

Keywords: Tonal Harmony, Augmented Reality, Physical Interaction, Mobile Devices.

Abstract: This paper presents *AREmbody*, an augmented-reality mobile application for the development of tonal harmony awareness. Continuing from previous prototypes based on full body and desktop interaction, *AREmbody* benefits from a very simple portable setup which allows physical interaction and supports the activity of one or more users. The application combines in a single mobile device a video processor, a media player and a movement tracker, opening the way to the design of harmonic games with challenges and recordable scores. Thus the application not only fosters music education activities in the classroom, but also extends them outside the school times and places, promoting educational practices in informal and private contexts.

## 1 INTRODUCTION


Augmented reality (AR) and virtual reality (VR) are attracting growing interest in the world of application designers, technological entrepreneurship and culture innovators (Becker et al., 2017).


The possibility offered by AR to create responsive environments where information may be superimposed to real objects is extremely appealing in many fields such as engineering analysis (Li et al., 2017), maintenance (Azuma, 1997), medical diagnosis (Douglas et al., 2017) and many others. On the other hand, VR offers further opportunities because it allows the creation of artificial responsive environments in which users are completely immersed and where they can manipulate virtual objects and obtain feedback in real time (Brooks, 1999). This has given life to dozens of applications in the fields of vehicle simulation (Kljuno and Williams, 2008), training (González et al., 2017) and healthcare (Moline, 1997). Even if affected by several limitations and poor dissemination due to technical difficulties


or high costs, AR and VR applications are gaining ground in the fields of architecture, engineering and construction (Chi et al., 2013) as well as through game-based approaches (Dinis et al., 2017).


Fred Brooks, one of the fathers of computer science, considered virtual reality as a medium primarily devoted to augment human intelligence and therefore particularly suitable for learning (Brooks Jr, 1996). This is the reason why virtual and augmented reality attracts the attention of educators and pedagogues, who try to develop efficient instructional approaches in order to take full advantage of the great potential offered by these technologies. Indeed, when shifting from practical and entertainment uses to education, a considerable amount of human factors issues emerge such as teachers' practices and training, different learning styles of the students, and the technological and didactic effectiveness of the applications (Keenaghan and Horváth, 2014).

Probably, the alignment of instructional design, systems' affordances, and didactic goals is the best way to exploit the learning potentials of AR and VR environments. Particularly, some interesting features emerge from AR and VR review analysis. One is the possibility offered by AR and VR of visualizing and interacting with abstract objects, such as the disposition of the planets (Lindgren et al., 2016) or the

<sup>a</sup> <https://orcid.org/0000-0002-1257-5878>

<sup>b</sup> <https://orcid.org/0000-0001-8435-8373>

<sup>c</sup> <https://orcid.org/0000-0002-8251-2231>

<sup>d</sup> <https://orcid.org/0000-0003-1863-4178>

combinations of chemical elements (Cai et al., 2014), which are not visible in real life. The opportunity of disclosing hidden concepts and processes makes them available for students' understanding and deeper knowledge (Wu et al., 2013). A second characteristic is spatial learning and function recognition, which implicitly derive from the process of visualization and are deeply involved when dealing with geometrical shapes (Kirner et al., 2012) or mechanical machinery (Poh et al., 2005). A third element is physical or full-body interaction, which is considered a strong means to create involvement and to foster learning processes (Price and Rogers, 2004).

The rest of the paper is organized as follows: in Section 2 we will present related work dealing with two aspects, namely the use of AR in music education and computer-based approaches to learn tonal harmony, which is the focus of our proposal; in Section 3 we will report the results achieved by previous experimentation that we conducted in the field; in Section 4 we will provide details about the AR-based app for iOS that we designed and developed with the aim to foster music harmony awareness in young students; finally, in Section 5 we will draw conclusions.

## 2 RELATED WORK

While the use of VR in music education has been investigated in various projects, we will focus on AR, that nowadays is a relevant field of interest and is more related to our work. The aim of many AR applications is supporting people in learning how to play musical instruments. Performing instructions for piano players – such as the keys to be pressed, fingering suggestions or the display of chords to be added to melodies – are projected on the keyboard through the use of a computer monitor (Barakonyi and Schmalstieg, 2005; Goodwin and Green, 2013), head-mounted displays (Chow et al., 2013; Fernandez et al., 2016; Hackl and Anthes, 2017), or ceiling-mounted projector (Rogers et al., 2014). There are also noticeable examples of applications for the study of other instruments, such as the acoustic guitar (Löchtefeld et al., 2011) or the Chinese guqin<sup>1</sup> (Zhang et al., 2015).

Likewise, music perception and music notation skills may benefit of AR applications. In (Lemos et al., 2017), a melody composed of painted notes is filled by the children on paper and then recognized in real time by a mobile's camera; once the image is acquired and processed, the notes may be played

<sup>1</sup>The Chinese guqin belongs to the family of plucked string instruments called citharas or table psalteries.

and checked by the system. A similar application, called *Augmented Songbook*, is based on pre-printed music sheets (Rusiñol et al., 2018). These gameful approaches are very engaging for children and, consequently, represent useful tools for music education. The simplicity and ubiquity of mobile devices make them suitable for use in the classroom as well as in other private contexts, thus fostering the diffusion of educational practices.

On the side of computer applications for the study of tonal harmony, a very traditional approach – that does not involve VR or AR approaches – may be found in the *MacGAMUT Music Software*, where exercises of tonal harmony dictation are organized in unities of growing difficulty<sup>2</sup>. The technical information is provided by the popular book of Koska and Payne about tonal harmony (Koska and Payne, 2013) for which MacGAMUT acts as a computer-based complementary worksheet.

A less academic and more playful approach can be found in the *Theta Music Trainer* platform, where the user can also find music games for tonic chord recognition<sup>3</sup> and for the identification of common chord patterns<sup>4</sup>. The commercial software *Mapping Tonal Harmony*<sup>5</sup> offers an interactive tool for visualizing a map of the tonal functions created from music analysis, while the *Chordify* platform<sup>6</sup> performs a song harmonization outlining chord names, chord changes, and instrumental positions and fingering for their performance. However, it must be highlighted that these useful tools are based on a high degree of the user's abilities, such as reading music notation, playing a harmonic instrument, or understanding chord notation and music structure.

## 3 BACKGROUND RESEARCH

The current proposal is rooted in previous research concerning the exploration of tonal music harmony and the development of harmonic awareness for children, students or non professional musicians.

The difficulty of explaining the complex theory behind tonal harmony and the need to avoid music notation led us to the use of spatial representations of the music structure. These offer the possibility of

<sup>2</sup><https://www.macgamut.com/>

<sup>3</sup><https://trainer.thetamusic.com/en/content/html5-tonic-finder>

<sup>4</sup><https://trainer.thetamusic.com/en/content/html5-speaker-chords>

<sup>5</sup><https://mdecks.com/>

<sup>6</sup><https://chordify.net/>

exploring the sound of the various chords and coupling them with precise locations in space. Thus, the timed harmonic changes occurring during a music piece produce precise movement patterns which can be easily learned and reproduced by the users. Once learned through embodied activities, movements can be employed in a computer environment for performance (Mandanici et al., 2019).

Technology has played a fundamental role in supporting these goals, by providing automatic tools for the detection of movements in the space (e.g., user's actions captured by a webcam), the performance of suitable music objects (e.g., chords triggered by specific events), and the collection of results (e.g., temporized sequences of chords saved in a database).

Concerning background research, it is worth mentioning two fundamental stages, which implied the design and release of computer-based solutions already tested in educational environments: *Harmonic Walk* and *Harmonic Touch*. These applications will be now described in detail.

### 3.1 Harmonic Walk

A first attempt to apply embodiment to music harmony awareness led to the design of a HW/SW system called *Harmonic Walk*, whose goal was to associate chord sequences to the movements of a user in a large-scale bi-dimensional environment thanks to spatial exploration and learning (Mandanici et al., 2016).

The user's interface consisted of a rectangular carpet positioned on the floor and reproducing a simplified version of a major tonality harmonic space. The visual reference for the location of the various chords was provided via specific cue points. The interaction happened when the user stepped on one of the interactive landmarks linked to the chord sounds.

The architecture of the system was roughly composed of two software modules, aimed at video analysis and sound synthesis respectively. A video camera mounted on the ceiling and oriented perpendicular to the floor captured the user's movements inside a rectangular area, whose dimensions depended from both from the camera-to-floor distance and the lens' field of view. As the system was designed for classroom activities, the goal was to monitor a limited area, about 3m×4m. The video module analyzed the images provided by the camera in order to detect the user's position. Coordinates were sent via OSC<sup>7</sup> to

<sup>7</sup>Open Sound Control (OSC) is a protocol for networking sound and multimedia devices for purposes such as musical performance or show control. It was originally intended for sharing music performance data, such as gestures, parameters and note sequences, between digital mu-



Figure 1: Session of *Harmonic Walk*.

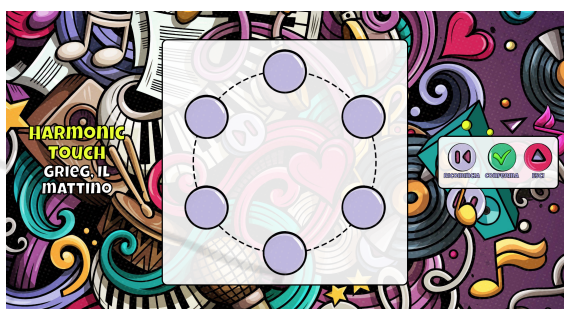


Figure 2: The interface of *Harmonic Touch*, exercise #1.

the sound synthesis module, implemented in the Max environment<sup>8</sup>.

Depending on the user's position, the sound module produced the sound of given chords. In this way, a free exploration of space let the player acquire awareness of music chords and learn the paths connected to the most common harmonic progressions.

### 3.2 Harmonic Touch

*Harmonic Touch* is a Web platform for the investigation and practice of tonal harmony. It has been conceived as a step-by-step path that leads users towards the discovery of important features of tonal harmony by leveraging on chord perception, gestural interaction and gamification techniques. Three groups of experiences are proposed to the user, focusing on: 1) the recognition of the implicit harmony, 2) the timed recognition of harmonic changes, and 3) melody harmonization. Each scenario has been discussed in detail in (Avanzini et al., 2019a).

*Harmonic Touch* recalls and extends the concepts of tonal-harmony space exploration introduced

sical instruments.

<sup>8</sup><https://cycling74.com/>

by *Harmonic Walk* by reviving them in a Web environment. Now, physical movements over a bi-dimensional surface can be performed on a PC screen, with no need to employ the HW/SW setup described in Section 3.1. The environment embeds both the on-screen movement tracker and the sound synthesis modules; moreover, the harmonic awareness achieved by students can be automatically assessed through a computer-based analysis of the data collected by the Web platform, as discussed in (Avanzini et al., 2019b).

In this context, we are mainly interested in the exercises dealing with the timed recognition of chord changes, which should constitute the final goal of the step-by-step process. Moreover, the AR application proposed in the following focuses on this scenario. Melody harmonization requires to select the right chords at the exact timing in order to accompany a music tune. The original piece, with the expected chord sequence, can be listened before the user starts playing. The activity consists in the recognition of the best-fitting chord among the proposed ones, whose selection should occur at a time as close as possible to the actual harmonic change. For this exercise, chords are represented along a circle, like in *Harmonic Walk*.

At present, the Web interface is undergoing a re-implementation mainly aiming to: 1) add a back-office section to manage music pieces, 2) obtain a better tracking of user's results, and 3) review the graphical user interface, so as to introduce responsive design and improve young users' engagement. The first and the second group of exercises are already available at <http://harmonictouch.lim.di.unimi.it/>, ready to be tested in primary schools after COVID-19 closure. The new interface is shown in Figure 2.

## 4 AREmbody: AN AR-ENHANCED APPROACH

The aim of the whole research project is to foster the development of music harmony awareness through embodiment. Recently, the idea of implementing a third application has emerged, on one side to simplify the setup required by *Harmonic Walk*, thus making the solution available for personal use, and, on the other, to improve the user's experience of *Harmonic Touch* by reintroducing movements in a physical space. To this goal, we have designed and implemented an AR-based application, currently available for iOS devices, called *AREmbody*. In this section, we will discuss technical aspects (Section 4.1), the game play (Section 4.2), and the advantages with respect to already developed solutions (Section 4.3).

### 4.1 Technical Remarks

*AREmbody* is a free app for iOS devices, mainly intended for iPhone and iPad. Due to the limitations imposed by Apple on the use of AR, it can be installed on models starting from iPhone 6s and running iOS 13 or a more recent version of the operating system. Currently, *AREmbody* is not available on the Apple App Store, but we are planning to release it in a public repository.

The app has been implemented in Swift, a programming language for macOS, iOS, watchOS, tvOS and other Apple platforms,<sup>9</sup>. The user interface has been realized in SwiftUI<sup>10</sup>.

The most technologically advanced part of the application adopts AR to let the user interact with the scene. In order to achieve this result, two frameworks by Apple, namely ARKit and SceneKit, have been used.

*ARKit*<sup>11</sup> aims to create a correspondence between virtual spaces and real spaces, thanks to the technique of visual-inertial odometry. This process combines the information coming from the sensors of the device (such as the accelerometer and gyroscope) with the analysis of the scene shot by the camera. In detail, the framework is able to recognize the most important characteristics of the scene, keep track of the differences in the positions of these cue points across frames provided by the camera, and compare this information with that coming from motion sensors. This combination allows the device to accurately model its position in space and its orientation. *ARKit* therefore connects the real world with the virtual one, i.e. the virtual space set by the developer to contain AR objects.

*SceneKit*<sup>12</sup> is a 3D graphics framework that lets the programmer create 3D scenes within an application. It combines a high-performance rendering engine with descriptive APIs that allow the import, manipulation and rendering of 3D assets. In order to structure the content of the scene, *SceneKit* implements a so-called *scene graph*, consisting of:

- the root node of the graph, which defines a coordinate space for the whole scene;
- other nodes that populate the scene, with visible content such as 3D assets.

The spatial and logical structure of a *SceneKit* scene is determined by the hierarchy of nodes that it contains.

<sup>9</sup><https://developer.apple.com/swift/>

<sup>10</sup><https://developer.apple.com/xcode/swiftui/>

<sup>11</sup><https://developer.apple.com/documentation/arkit/>

<sup>12</sup><https://developer.apple.com/documentation/scenekit/>

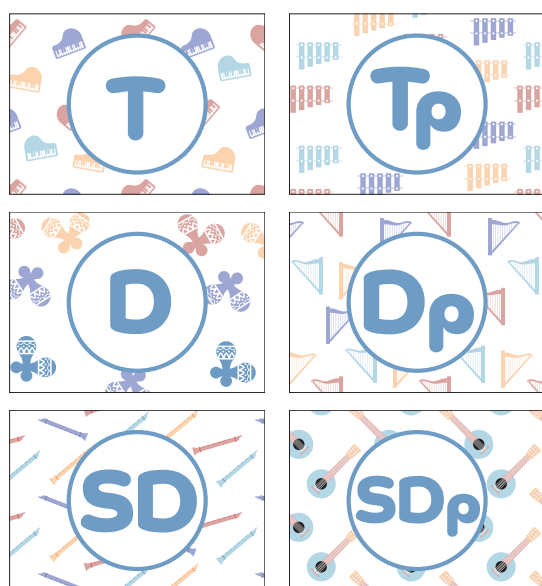


Figure 3: The six markers.

*AREmbody* integrates *ARKit* and *SceneKit* technologies in order to analyze the scene, recognize the presence of custom graphical signs called *markers*, and add the corresponding AR objects.

Since the goal is to foster harmony awareness, both markers and AR objects are related to tonal functions, and, more specifically, chord symbols. Following Riemann's harmonic theory (Riemann, 1896), the tonal space can be divided into:

- *Primary chords*, called tonic (T), subdominant (SD), and dominant (D), built on the I, IV, and V grade of the major scale respectively;
- *Parallel chords*, called parallel tonic (Tp), parallel subdominant (SDp), and parallel dominant (Dp), built on the VI, II, and III grade of the major scale respectively.

For further details about the music theoretical background, please refer to (Mandanici et al., 2019).

In order to improve their recognition by *ARKit*, the images representing chords should be as different as possible, like photos portraying completely inhomogeneous subjects; but, in this case, the goal was also to keep a graphical coherence among markers. As the result of a compromise, images are characterized by distinguishing chord letters and randomly disposed background symbols, as shown in Figure 3.

## 4.2 User's Interface and Game Play

*AREmbody* is an app that presents a number of views, including game instructions, contact, and a setup win-



Figure 4: Recognition of markers and superimposition of chord symbols in AR.

down. The most relevant view, representing the user's interface during the game play, is shown in Figure 4.

One of the noticeable aspects of *AREmbody* is the possibility to manage on a single device a number of different users, also organized by school and class. In this way, a game session addressing a group of participants reduces the dead time for setting up the app and switch players. Moreover, a single app instance can trace, analyze and compare different users' results across a game session, an activity spanning over different sessions, and even an experimentation conducted in different classes and schools. In the lower part of Figure 4, it is possible to notice the name of the current player (in this case, "Mauro") and the controls to switch to the previous/next player.

The other key aspect is the choice of the piece to harmonize, which happens before the game session starts. During the game play, the controls in the lower part of the interface allow to start and stop the playback. Moreover, a switch lets the player enter a sandbox, so as to listen to the reference piece: in this scenario, the original song is played with its correct harmonization, and user's actions are not tracked.

During the playback, the song is reduced to its

leading voice, whereas accompanying chords are demanded to user's actions. Specifically, the gesture that triggers the production of a given chord is the occlusion of the corresponding marker. As far as the marker is covered in the scene framed by the camera, the chord is considered *active*, and it is turned off as soon as the marker becomes visible again.

Different means can be used to enable/disable chord recognition, also depending on markers' size and layout: the whole body, a body part (e.g., hands or feet), a covering object (such as a paper sheet, a book, or a paddle), an interaction with markers (such as flipping the sheet). The system promptly reacts to these events and records the timing with reference to the song's position.

The AR component superimposes chord symbols over the corresponding markers, mainly with two goals: 1) making them more easily retrievable during the game play, and 2) confirming that the tracking system of the device is correctly managing their recognition. Concerning the latter aspect, AR symbols appear when markers are present in the scene, and disappear when they are no more recognized, due to framing (the marker is actually away) or to temporary occlusion (the player is hiding it). Further possible uses, not implemented in the current version of *AREmbody*, will be proposed in the next section.

At the end of each session, results are stored in the device. The app presents a function to export them in JSON,<sup>13</sup> so as to feed a database or an external analysis module. The non-trivial research problem of how to assess this kind of raw data has been investigated in (Avanzini et al., 2020).

### 4.3 Main Advantages

*AREmbody* is basically a portable and personal version of *Harmonic Walk*, focusing on the third group of exercises of *Harmonic Touch* and equipped with advanced features.

The first goal was to simplify the setup requirements of *Harmonic Walk*, but preserving the body interaction within a physical space. Nowadays, a solution is offered by mobile devices, which typically embed video cameras and currently have computational power sufficient to process images, track markers and produce sound. With respect to the HW/SW equipment required by *Harmonic Walk*, a mobile phone can be seen as an all-in-one device, acting as a video processor, a media player, a movement tracker, a multiple-user performance recorder, and a

<sup>13</sup>JSON, standing for JavaScript Object Notation, is a lightweight data-interchange format based on structured plain text.

local database with export functions. The app can run on a tripod-held as well as a hand-held device, as shown in Figures 5 and 6.

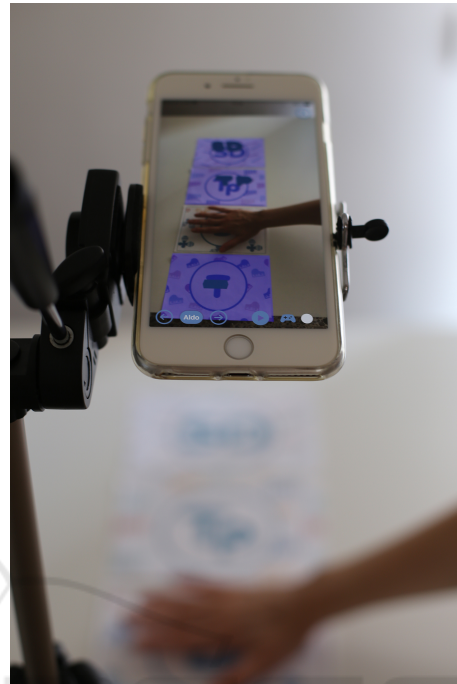


Figure 5: The app running on a tripod-held device.



Figure 6: The app running on a hand-held device.

*AREmbody* supports not only a personal, but also a personalized user's experience. Many aspects of the educational activity previously constrained by the setup can now be customized, ranging from the appearance of printable markers (graphical content, size, material, etc.) to their layout in the physical space (number of markers, position, expected body interaction, etc.). These features introduce flexibility in the educational experience, which can take into account visual, physical and cognitive impairments. Concerning the first aspect, it is possible to emphasize graphical differences between chord signs by stressing color contrast, increasing markers' sizes, using bigger or more readable fonts, etc. Moreover, visual markers can be associated with tangible cues for BVI<sup>14</sup> people. With respect to physical impairments, users are not required to walk on a carpet (which is still a possibility), but, for instance, they can also occlude markers on a table by hand or move the camera away: these gestures would equally preserve the relationship between space and harmony. Finally, when considering cognitive impairments, the exercise can be suitably simplified, for example by reducing the number of choices or adopting other spatial layouts (e.g., a linear instead of a circular one). Even if the *AREmbody* game play can be user-tailored, the mentioned choices should be supervised by an expert, so that the pedagogical efficacy is preserved.

With respect to its potential, the AR component of the game play is probably underused. In the current version of *AREmbody*, the main purpose of AR is to enrich the user's experience, basically to highlight the recognition of markers, to reinforce their position, and to make the game play more appealing. In a future version, AR could better support a number of user-defined features. First, colors, font types, text dimensions, etc. could fall under the user's control, so as to improve the user's experience, also with the side effect of helping impaired players. Besides, graphical aspects could adapt to specific environmental scenarios, such as poor lighting conditions or crowded scenes. Finally, the AR content itself could become more informative, providing alternative text or guidance in form of animations. For example, a visual hint could indicate the proximity of a chord change and/or suggest the right choice.

## 5 CONCLUSIONS

In this paper we have presented *AREmbody*, an iOS app aiming to foster tonal harmony awareness in

<sup>14</sup>BVI stands for Blind or Visually Impaired.

young students through embodiment. The capability to analyze the scene captured by the camera and to add AR objects to its visualization in the display provides a gamified approach to interact with the real environment, which is expected to produce better learning results. The app also provides a number of features to improve performance assessment, ranging from the support to articulated game play sessions (multi-player, multi-class, etc.) to the possibility to collect, export, and finally analyze results.

Concerning future work, we are mainly exploring three directions. First, we want to address a wider audience of users: since the iOS app requires relatively recent and expensive devices, we are studying the feasibility of an Android version, that should run also on low-budget smartphones. A second goal is to extend the functionalities in order to cover all the exercise types available in *Harmonic Touch*, that responded to a well-defined pedagogical path: recognition of the implicit harmony, timed recognition of harmonic changes, and melody harmonization; conversely, the current release of *AREmbody* focuses on the final step only. Finally, in September 2020 we expect that school activities can restart after the lockdown due to COVID-19 emergency. This would allow to test the educational achievements of *AREmbody* in a real-world scenario.

## REFERENCES

- Avanzini, F., Baratè, A., Ludovico, L. A., and Mandanici, M. (2020). A web platform to foster and assess tonal harmony awareness. In Lane, H., Zvacek, S., and Uhomobhi, J., editors, *Computer Supported Education – 11th International Conference, CSEDU 2019, Heraklion, Crete, Greece, May 2-4, 2019, Revised Selected Papers*, volume 1220 of *Communications in Computer and Information Science*. Springer International Publishing.
- Avanzini, F., Baratè, A., Ludovico, L. A., and Mandanici, M. (2019a). A computer-based approach to teach tonal harmony to young students. In Lane, H., Uhomobhi, J., and Zvacek, S., editors, *Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019)*, volume 1, pages 271–279. SCITEPRESS - Science and Technology Publications, Lda.
- Avanzini, F., Baratè, A., Ludovico, L. A., and Mandanici, M. (2019b). Metrics for the automatic assessment of music harmony awareness in children. In Barbancho, A. M., Barbancho, I., Peinado, A., and Tardón, L. J., editors, *Proceedings of the 16th Sound & Music Computing Conference (SMC 2019)*, pages 372–379. SMC.
- Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385.

- Barakonyi, I. and Schmalstieg, D. (2005). Augmented reality agents in the development pipeline of computer entertainment. In *International Conference on Entertainment Computing*, pages 345–356. Springer.
- Becker, S. A., Cummins, M., Davis, A., Freeman, A., Hall, C. G., and Ananthanarayanan, V. (2017). NMC horizon report: 2017 higher education edition. Technical report, The New Media Consortium.
- Brooks, F. P. (1999). What's real about virtual reality? *IEEE Computer graphics and applications*, 19(6):16–27.
- Brooks Jr, F. P. (1996). The computer scientist as toolsmith II. *Communications of the ACM*, 39(3):61–68.
- Cai, S., Wang, X., and Chiang, F.-K. (2014). A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior*, 37:31–40.
- Chi, H.-L., Kang, S.-C., and Wang, X. (2013). Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in construction*, 33:116–122.
- Chow, J., Feng, H., Amor, R., and Wünsche, B. C. (2013). Music education using augmented reality with a head mounted display. In *Proceedings of the Fourteenth Australasian User Interface Conference-Volume 139*, pages 73–79.
- Dinis, F. M., Guimarães, A. S., Carvalho, B. R., and Martins, J. P. P. (2017). Virtual and augmented reality game-based applications to civil engineering education. In *Global Engineering Education Conference (EDUCON), 2017 IEEE*, pages 1683–1688. IEEE.
- Douglas, D. B., Wilke, C. A., Gibson, J. D., Boone, J. M., and Wintermark, M. (2017). Augmented reality: Advances in diagnostic imaging. *Multimodal Technologies and Interaction*, 1(4):29.
- Fernandez, C. A. T., Paliyawan, P., Yin, C. C., and Thawonmas, R. (2016). Piano learning application with feedback provided by an ar virtual character. In *2016 IEEE 5th Global Conference on Consumer Electronics*, pages 1–2. IEEE.
- González, A. V., Kapalo, K., Koh, S. L., and LaViola, J. J. (2017). Exploring the virtuality continuum for complex rule-set education in the context of soccer rule comprehension. *Multimodal Technologies and Interaction*, 1(4):30.
- Goodwin, A. and Green, R. (2013). Key detection for a virtual piano teacher. In *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, pages 282–287. IEEE.
- Hackl, D. and Anthes, C. (2017). HoloKeys – an augmented reality application for learning the piano. In *Forum Media Technology*, pages 140–144.
- Keenaghan, G. and Horváth, I. (2014). State of the art of using virtual reality technologies in built environment education. *Engineering: Education and Innovation*.
- Kirner, T. G., Reis, F. M. V., and Kirner, C. (2012). Development of an interactive book with augmented reality for teaching and learning geometric shapes. In *Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on*, pages 1–6. IEEE.
- Kljuno, E. and Williams, R. L. (2008). Vehicle simulation system: controls and virtual-reality-based dynamics simulation. *Journal of Intelligent and Robotic Systems*, 52(1):79–99.
- Kostka, S. and Payne, D. (2013). *Tonal harmony*. McGraw-Hill Higher Education.
- Lemos, B., Correa, A., Nascimento, M., and Lopes, R. (2017). Augmented reality musical app to support children's musical education. *Journal of Computer Science and Information Technology*, 5(4):121–127.
- Li, W., Nee, A., and Ong, S. (2017). A state-of-the-art review of augmented reality in engineering analysis and simulation. *Multimodal Technologies and Interaction*, 1(3):17.
- Lindgren, R., Tscholl, M., Wang, S., and Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, 95:174–187.
- Löchtefeld, M., Gehring, S., Jung, R., and Krüger, A. (2011). guitarAR: supporting guitar learning through mobile projection. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1447–1452.
- Mandanici, M., Ludovico, L. A., Avanzini, F., and Baratè, A. (2019). Learning tonal harmony through bodily interactions and gamification. In Arculus, C., Nijs, L., and Van Regenmortel, H., editors, *MERYC19. Counterpoints of the senses. Bodily experiences in music learning*, pages 237–246. Ghent University.
- Mandanici, M., Rodà, A., and Canazza, S. (2016). The Harmonic Walk: An interactive physical environment to learn tonal melody accompaniment. *Advances in Multimedia*, 2016.
- Moline, J. (1997). Virtual reality for health care: a survey. *Studies in health technology and informatics*, pages 3–34.
- Poh, Y., Nee, A. Y., Youcef-Toumi, K., and Ong, S. (2005). Facilitating mechanical design with augmented reality. *Innovation in Manufacturing Systems and Technology (IMST)*.
- Price, S. and Rogers, Y. (2004). Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education*, 43(1-2):137–151.
- Riemann, H. (1896). *Harmony simplified, or the theory of the tonal functions of chords*. Augener Ltd.
- Rogers, K., Röhlig, A., Weing, M., Gugenheimer, J., Könings, B., Klepsch, M., Schaub, F., Rukzio, E., Seufert, T., and Weber, M. (2014). Piano: Faster piano learning with interactive projection. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, pages 149–158.
- Rusiñol, M., Chazalon, J., and Diaz-Chito, K. (2018). Augmented songbook: an augmented reality educational application for raising music awareness. *Multimedia Tools and Applications*, 77(11):13773–13798.
- Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., and Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & education*, 62:41–49.
- Zhang, Y., Liu, S., Tao, L., Yu, C., Shi, Y., and Xu, Y. (2015). Chinar: facilitating Chinese Guqin learning through interactive projected augmentation. In *Proceedings of the Third International Symposium of Chinese CHI*, pages 23–31.