



Collaborative Music Composition Based on Sonic Interaction Design

Mauro Amazonas¹, Victor Vasconcelos¹, Adriano Brandão¹, Gustavo Kienem¹,
Thaís Castro^{1(✉)}, Bruno Gadelha¹, and Hugo Fuks²

¹ Federal University of Amazonas (UFAM), Manaus, AM, Brazil
{mauro.jr,vfv,thais,bruno}@icomp.ufam.edu.br,
abgroovy2@yahoo.com, gustavo_gustavo1@hotmail.com

² PUC-Rio, Rio de Janeiro, RJ, Brazil
hugo@inf.puc-rio.br

Abstract. This work proposes a way to create sonic interaction through collaborative composition of spatialized sounds in real time, based on SID (sound interaction design). There is no centralizing figure for the sound designer, and each person may begin a sonic interaction spontaneous and independently, or join an ongoing interaction. Therefore, each person is responsible for the result of his own interaction. Every movement a person makes within the reach of the determined space, the sound is captured and processed and then externalized as feedback in the multichannel sound system, perceived as a unified sound. The concepts and app designed for sonic interaction discussed in this work is intended for use in multidisciplinary contexts, raising important technical challenges. We have devised a design process resulting in four different prototypes, attending to different perspectives. Each prototype had its own experiment.

Keywords: Collaborative music composition · Real time spatialization
Interaction design

1 Introduction

Performing music is a natural collaborative activity. Although there is the possibility of a song being performed by a single interpreter (i.e. singer, guitarist or violinist), usually music performances are within groups, bands or orchestras, where each musician collaborates playing with his instrument and the audience have the perception of hearing all of them as one unified sound. This perception is achieved by the effective communication among different instruments and musicians, coordinated by the maestro.

Technology have been promoted new ways of collaboration regarding music creation and performance. Especially, electroacoustic music, one of the many music facets, dives deeply in the potentialities provided by technology. Since its rising in the end of the years 1940, electroacoustic music shows an intimacy with technology, acousmatic, mixed or derived from electronics in real time, such as Live Electronics [13]. An electroacoustic product is composed by sounds, previously recorded or live performed, later

treated in a computer and refined musically in order to be part of the final product, may have synthesized sounds or from nature [5].

The artists have been combined emerging technologies with art to produce new music expressions [2]. Simple objects as balls, floaters, microphones and sensors become means for interaction, each of them performing its own characteristic sound, resulting in a composition [8]. Other use of technology in playful, artistic and creative experiences, is obtained when mixing art with technology. Based on Sonic Interaction Design (SID) concepts, some authors [10] have developed interactive sound interfaces using computing vision software, named ReactiVision. According to [12], SID is the study of sound exploration as a means of information exchange, meaning and emotional aspects in interaction. As a result, SID belongs to both domains: interaction design and computational music.

There are also other uses for SID. For example, in movie industry. SID is behind the creation of immersive effects through multiple sound sources. People barely noticed that there is a specific design for the sound and this is the way the sound is presented, involving the audience in a spatialized sound field [4]. Other example of SID, and may be the oldest use, is in videogames.

This work proposes a way of creating sonic interaction through collaborative spatialized sound composition in real time, based on SID. In our proposal, there is not a centralized role for the sound designer, meaning that each person who takes part of the interaction may initiate a spontaneous and independent sonic interaction or may join an ongoing one. Consequently, each person is responsible for the result of his own interaction. In practical terms, every movement a person makes towards the determined space is captured, processed and externalized as a feedback in the multichannel sound system, perceived as a unified sound.

This work's scope relies on electroacoustic music, where collaboration become art as an original source for music composition. Many people play a part in interaction space with their own sounds, bringing up the style and random principle Pierre Boulez [7] to result in music and random collaborative music. All those interaction is processed and spatialized by the system in real time.

2 Method

The spatialization environment for sonic interaction design described in this work was achieved through the evolutionary prototyping approach. In this approach, each prototype is evaluated at a time, having different goals [3]. Altogether, the prototypes are part of the product prototype.

The process of evolutionary prototyping, usually, follow four iterative steps: (i) identification of user basic needs; (ii) development of a working prototype; (iii) implementation and use of the prototype; (iv) revision and enhancement. In the case of this work, this method was chosen for having these different possibilities for testing different technologies and yet result in a product prototype. This was, prototypes are extremely useful in the discussion with the users, being a communication device among members

of the development team and an effective way for the designers exploring their design ideas [13].

The product prototype resulted from the evolutionary process was built combining and testing concepts of four different prototypes, according to distinct goals: (1) testing the initial composition idea using simulators, through the simultaneous execution of user sounds which enter and leave the space; (2) same technology behind, but adjustments were made for testing with real users, in real time; (3) using location-based features for detecting users in the space and directing their sounds to the closest speaker; and (4) simulating geolocation with controls for using the prototype in closed spaces.

3 Prototyping for Collaborative Sonic Interaction

According to the evolutionary prototyping approach, as aforementioned, we developed four prototypes for testing different features. Although each one uses a specific technology, they all share resources for spatialized sonic interaction, as an amplified sound system, a delimited space where the interaction happens and a device for initiating the interaction.

The focus of the proposed environment is to foster interaction among people using the sound each person chosen to play and represent herself. Interaction space can be any room or an open space with a delimited area. Speakers must reproduce sounds from people who are intentionally in the interaction delimited space. Once a person leaves the space, the system turns off her corresponding sound automatically. People may enter and leave the interaction space, tuning on and off their sounds, activating and deactivating the speakers in real time. Other feature for the proposed environment is that's when moving around, people are represented by their sounds and the direction of that sounds moves with each person. Furthermore, every composition executed at any moment with individual sounds is recorded and reproduced in real time.

In order to give appropriate support for the aforementioned scenario, some computational resources are needed. The overall architecture of the product prototype is described through a schematic diagram, in Fig. 1, showing which and how those resources are used as well as the environment's dynamic.

According to Fig. 1, there is a server with double functionality, acting as a web server for storing user profiles and an audio server, for processing real time spatialization of users' chosen sounds. Four amplified speakers are connected to the audio server for giving immediate feedback to the interaction space. Router A, which works as an access point, allows the connection between a client app and a server app. Router B, which does not have any kind of connection with the client and server works only as a central point, delimiting the interaction space. This router B radius defines whether a person is within the reach of the interaction space. Whenever the client app is able of detecting router B signal, the sound is reproduced. Otherwise, the sound is interrupted.

Following we present and discuss this environment's evolutionary process, main goals, experimentation and analyses for each one of them.

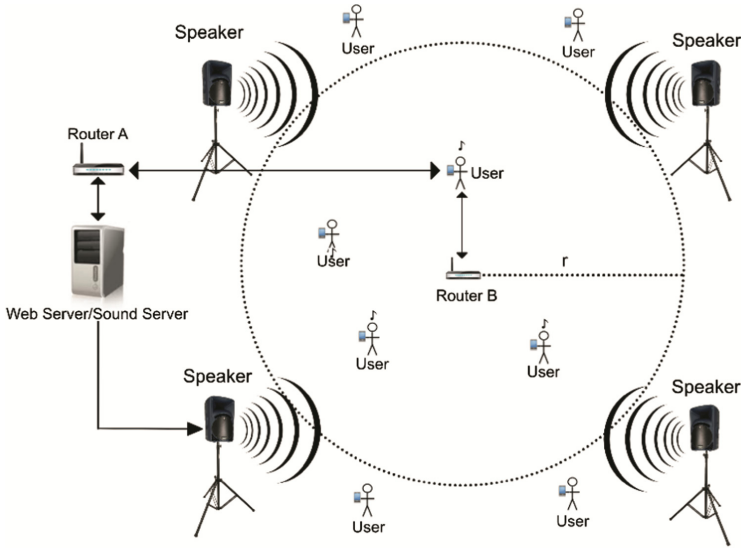


Fig. 1. System's schematic diagram.

3.1 Prototype 1 – Viability

In order to evaluate the viability of an implementation from the server side using Pure-Data [11] it was necessary to develop a concept prototype and carry on a study to check whether PureData were able to deal with many users reproducing different sounds at the same time. PureData proved to be robust enough to deal with many users and sounds at the same time.

Given that this study had a technical objective and its results would guide this research specific technology adoption, it wasn't necessary to carry out an experiment with users. However, the prototype simulates the interaction space using a simulation software, named TUIO Simulator [15], represented in Fig. 2.

TUIO was adapted for giving support to the most quantity of geometric figures as possible. Each geometric figure represents a person in the interaction space. In addition to that, they can "walk" on the circle in the middle, selecting one of the twenty possible sounds.

The circle in the middle represents the interaction space. When a figure is inside the circle a message is sent to the server's app and the corresponding sound starts to play. When the user leaves the interaction space (removing the figure) other message is sent to the server's app which finishes the sound reproduction.

As the simulation pointed out, it proved viable to adopt PureData as the environment platform. This conclusion came after the successful reproduction (with no overload) of many sounds at the same time. The platform was robust enough to support many possibilities and different quantity of requests.

In this prototype, all users (people) were modeled each one at a time, with the upper limit of fifty. In this way, if it was necessary that more users used the app, each one could be added manually. Although being flexible regarding the number of users, this kind of

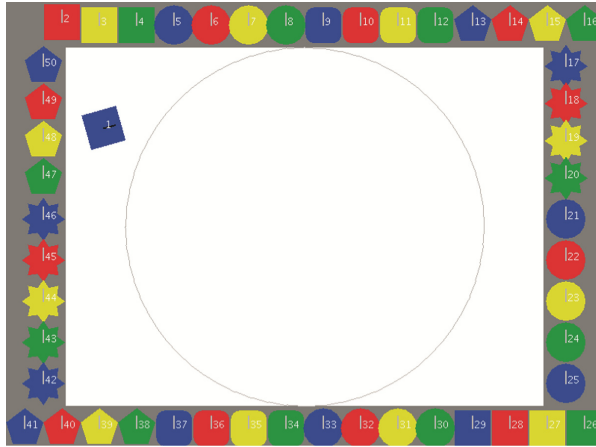


Fig. 2. Virtual interaction on TUIO simulator.

implementation had some flaws, as the one about code's correction. If there were implementation errors, it would be hard to find them because the number of elements on screen was considerably big and complex once PureData has difficulties in dealing with more than 50 objects.

Other critic aspect of this prototype was the use of the TUIO communication protocol. This protocol has a particular syntax to stablish communication between device and server. When using the simulator, the protocol works just fine, but when using it with a mobile device, some problems appeared. As a result, we decided to use the Open Sound Control (OSC) protocol for the next prototype. One instance of this test is available at the internet address: [1].

3.2 Prototype 2 – Mobile

Prototype 2 involved a client application developed for mobile devices for Android. This application is represented in Fig. 3. In this case, an empirical study was conducted with real users, yet without sonic spatialization, for identifying users who were inside or outside the interaction space.

The mobile app, named Compomus (from MUSic COMPosition), was developed only for Android because the potential users are the educational community from the University of Amazonas, in which almost 80% has a smartphone with Android.

Compomus has three main screens (Fig. 3). Screen (A) is for user's information display, where it's mandatory for the user to register a user name, email and password. Then, it shows screen (B), where the user chooses one of the sounds already stored on the database. It's worth mentioning this sound represents the user in the interaction space. All the sounds are from birds of the Amazon Forest. These samples were used to try recreating the sonic scene of the forest.

Next screen is (C), where the app informs user's id, his status (inside or outside the interaction space) and a button for sound change, which can be pressed anytime.

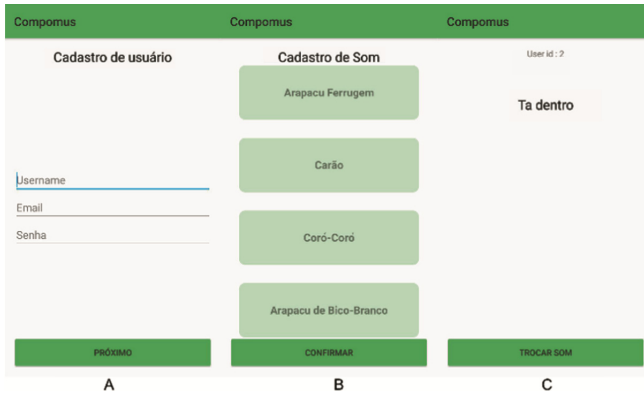


Fig. 3. Prototype 2: compomus.

Empirical Study. According to [6], viability studies must be conducted when using different technologies. For testing the application on both sides, client and server, we decided to recruit real potential users. Still during planning, the questions raised were (i) participants will notice their interaction in the interaction space?; (ii) participants will notice audio feedback from the speakers?; (iii) participants will notice the collaborative composition in real time? Following we present empirical study’s elements.

Participants. Ten students enrolled in the Collaborative Systems course volunteered to participate. They were already familiarized with the concepts involved in the use of the prototype but they did not have any knowledge or previous experience with music composition, production or theory.

Location. A pre-defined space in a classroom was reserved for the test. For doing that, the router responsible for propagating the sounds in the middle of the interaction space was configured for this limitation. Only one amplified speaker was used, located in the middle of the interaction space.

Tasks. Participants had six tasks: (1) to install Compomus; (2) To fill the registration form with their information; (3) To select a sound to represent them; (4) To interact freely, in and out the interaction space; (5) To change their sound; and (6) To complete a post-test questionnaire. After task 3, all students had 15 min to complete tasks 4 and 5. After that, they completed task 6.

The three questions (i, ii, iii) that guided this study are about people’s awareness, regarding audio, collaboration and interaction. Regarding their overall experience, analyzing the post-test questionnaire, we surprisingly found out that all of them deemed important their participation in the global composition during interaction, being able to have all awareness elements, raised by the questions.

Although their overall experience was good using Compomus, they experienced also some difficulties during their experience. One example was the lack of graphical feedback because they had to rely only on the sound and the written status. At some point, a student’s status wasn’t refreshed, he didn’t hear his sound and became confused.

One explanation for the synchronization problem reported above is that we noticed a ten seconds delay when Android system look for new networks. The same problem happened when a student left the interaction space and his sound continued to be executed for some time.

With that study, we noticed that the resultant composition of sounds (music) is always different, even when using the same scenario configuration, what puts in evidence the collaborative aspect of this kind of interaction. Other aspect is that besides participants describe their experience as a good one, their interaction was limited only to their entrance and exit. This aspect had to be improved to foster more interaction.

3.3 Prototype 3 – Four Channels Spatialization

Considering the results from the study of prototype 2, server application was totally re-implemented, with a more robust code to increase performance and to become more scalable. It consists of techniques to generate executable codes, sub patches and dynamic patches, solution that drastically decreased the number of graphical elements in screen, as shown in Fig. 4. Besides that, it was added a sonic spatialization support for four channels, using a cross-art quadraphonic panorama technique.

In Compomus, it was implemented a location based feature using GPS technology to automatically direct users' sounds to a speaker. As each user uses the app on his smartphone, data from his smartphone, the app request the data and sends it to the audio server, which interpret the information, performing the real time spatialization.

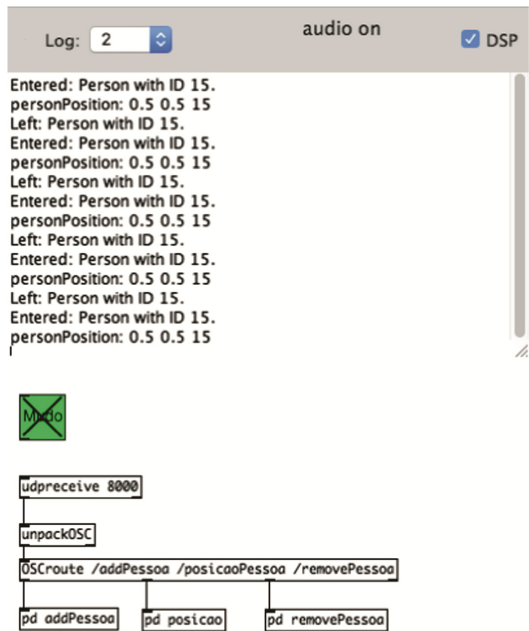


Fig. 4. Prototype 3: server app.

Empirical Study. Within this new solution, another question was raised: Will the participants notice their own sound being spatialized according to their location in the interactive space? This study consisted of the same elements as the previous one, adding only the complexity of spatialization for the user's point of view and a change in the location. By this means, it was set as follows. Figure 6 shows the test.

Participants. Ten students enrolled in the Collaborative Systems course volunteered to participate. They were already familiarized with the concepts involved in the use of the prototype but they did not have any knowledge or previous experience with music composition, production or theory.

Location. A pre-defined space (for technical restrictions as the size of the cables) in an open space, in a hall of the University of Amazonas. For doing that, the router responsible for propagating the sounds in the middle of the interaction space was configured for this limitation. Four amplified speakers were used, distributed on the corners of the interaction space.

Tasks. Participants had four tasks: (1) To select a sound to represent them; (2) To interact freely, in and out the interaction space; (3) To change their sound; and (4) To complete a post-test questionnaire. All students had 15 min to complete tasks 1, 2 and 3. After that, they completed task 4.

This test was very relevant to the research because it confirmed our assumptions about the imprecision of GPS in smartphones. System's overall performance in the location wasn't as good as it was expected because due to technical limitations, the space was very restricted and for having an upper floor, making it difficult to get more precision from the GPS.

Participants said they could spatialized their sounds and their location correctly, even for a short time. However, many GPS imprecisions were also noticed by the participants, as jumps in location and errors. For instance, a participant was in a spot and the GPS indicated he was on the opposite side. Due to these difficulties, with this technology, the interaction space has to be in a more open space at the University, as in the student hall in the campus entrance (Fig. 5).

Besides problems with GPS students said they had a good experience with spatialization and the felt they were immersive in the space because of the number of speakers and the change in the source of the audio feedback. They also felt powerful for being able to control the source of feedback for their sound by walking around the interaction space.

3.4 Prototype 4 – Closed Spaces

As the client app was well accepted by participants, the fourth prototype consisted mainly on the server's side, including performance enhancements and the swap of the cross-art quadraphonic panorama technique for the Ambisonics [9] technique, that provides more flexibility regarding the increase of speakers and also precision in the virtualization of the sonic source, using the open access library Open Audience.



Fig. 5. Spatialization test with people.

Compomus, was configured in such a way that it could test the new auralization technique and improving sonic spatialization. This way, a new screen (C) was added, as shown in Fig. 6. Screen (C) has a joystick where the participant can direct his sound to wherever he wants as he was walking in the interaction space.

Empirical Study. Different from the previous studies, this one had four new volunteers, non-familiarized with the environment for collaborative sonic interaction, shown in Fig. 8.

Participants. Two teachers and two students volunteered to participate. They were already familiarized with the concepts involved in the use of the prototype but they did not have any knowledge or previous experience with music composition or production. One of them had experience with music theory.

Location. A pre-defined space in an acoustic room, in the Arts Center at the University of Amazonas. For doing that, the router responsible for propagating the sounds in the middle of the interaction space was configured for this limitation. Four amplified speakers were used, distributed on the corners of the interaction space.

Tasks. Participants had four tasks: (1) To select a sound to represent them; (2) To change directions in the interaction space; (3) To identify his own sound among others; and (4) To complete a post-test questionnaire. All students had 15 min to complete tasks 1, 2 and 3. After that, they completed task 4.

Considering the interaction space, a small area in a rehearsal studio was prepared. In this case, router (A), that would delimit the interaction space, wasn't necessary



Fig. 6. Compomus – screen for spatialization.

because participants had the possibilities of coming in and going out of the system whenever they wanted.

As aforementioned, this prototype focus is the spatialization using the Ambisonic auralization technique and the sound spatialization using the library Open AUDIENCE. This empirical study is more about improving the sonic experience than the application itself. For that reason, the test took place in a more confined space and the tasks were slightly different from the ones with the other prototypes (Fig 7).



Fig. 7. Test with prototype 4.

The last task, common to all prototypes, was a post-test questionnaire. This time, participants answered seven questions about their experience with the space, interaction, collaboration and spatialization. One of the participants had some difficulty in the beginning of the interaction using Compomus, due to his smartphone has frozen a few times. Other participants had a good experience with Compomus and with the interaction space as a whole.

Regarding their awareness about the mixture of the different sounds they were aware of each sound and the composition made by each one. About the identification of his own sound, one participant noticed that whenever he directed his sound to one speaker and another participant direct his sound as well, the first sound stopped being performed. It did not happen with other participants.

All of the participants correctly identified their own sounds and with a good precision about the location and they could Interact with each other trying to guess where his sound was.

The mixture of sounds, according to participants statements, did not affect their awareness about their own sound and the other sounds individually. Besides a few problems with spatialization, they had a good experience with the interaction space and they enjoyed the possibility of creating some kind of music composed by individual sounds. In the way, this prototype was conceived, without using GPS technology, it also proved to be possible to take this spontaneous interaction experience to other spaces, where there is much interference with the GPS signal.

4 Conclusion

The concepts and applications projected for the sonic interaction environment discussed in this work resulted from a partnership between researchers from the Informatics Institute and Faculty of Arts, both from the University of Amazonas. This multidisciplinary context raises technical challenges for both areas. From computing point of view, there are challenges about how to provide interaction and make the process transparent for the user, involving sensors, network connections, interaction design, accessibility and user experience. From arts point of view, the challenges are in analyzing the resultant sounds, changing sound perspective from consumer to producer and dealing with sound production with sound production without total control.

In this work, we discussed how sonic spatialization can foster spontaneous collaboration in a pre-defined space, as a museum room, hall in a University campus or even open spaces. In order to get good results, we emphasize that the coworkers in different research areas may achieve interesting results. The spatialization using SID have already been explored in games, cinema, virtual reality, artistic performances and musical performances. The problem is that spatialization is usually pre-recorded in those contexts, what loses a potential for collaboration. For that reason, this work brings innovation by introducing more interaction possibilities for places where people only pass by.

From the evolutionary prototyping used in the research we have some possibilities for extensions and development of the product prototype. Possible improvements for the current environment are: improvements in redundancy of the information received;

studies about tridimensional configurations for speakers, according to the Octonal Cubic model; portability to a structured language as Super Collider, to make this system stand-alone, without the need of the PureData and with the possibility of being embedded in specific software, as indie games, VR apps. Possible improvements for a product prototype are: improvements in the graphical user interface according to interaction evaluation; development of new modules for the sonic spatialization system to get a more precise identification of the sound source; and to changing system's modules and components to be external functions for the PureData environment.

References

1. Amazonas, M.J.B.: <https://www.youtube.com/watch?v=913Y3yew5sw>. Accessed 01 Oct 2018
2. Blaine, T., Fels, S.: Contexts of collaborative musical experiences. In: Proceedings of the International Conference for New Instruments for Musical Expression (NIME-2003), pp. 129–134 (2003)
3. Carr, M., June V.: Prototyping and software development approaches. Department of Information Systems, City University of Hong Kong, Hong Kong, 319–338 (1997)
4. Costa, N.J.B.: O Surround e a Espacialidade Sonora no Cinema. Dissertação (Mestrado)—Escola de Belas Artes - UFMG (2013)
5. Fritsch, E.F.: Musica Eletrônica - Uma Introdução Ilustrada. Ed. da UFRGS (2008)
6. Shull, F., Carver, J., Travassos, G.: An empirical methodology for introducing software processes. In: CM SIGSOFT Software Engineering Notes (2000)
7. Guberniko, C.: Pierre Boulez e o pensamento musical da segunda metade do século XX: XVI Congresso da Associação Nacional de Pesquisa e pós-graduação em Música- ANPPOM, Brasília (2006)
8. Jensenius, A.R., Voldsund, A.: The music ball project: concept, design, development, performance. In: Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2012), pp. 300–303 (2012)
9. Lecomte, P., Gauthier, P.-A.: Real-time 3D ambisonics using faust, processing, pure data, and OSC. In: 15th International Conference on Digital Audio Effects (DAFx-15), Trondheim (2015)
10. Nas, E., Lopes, J.: Design de interfaces para experiências lúdicas com novas tecnologias: O caso do software-livre reactivision em interações sonoras. In: Blucher Design Proceedings, pp. 4960–4970 (2016)
11. Puckette, M.: Pure Data. <http://www.puredata.info>. Accessed 06 Jan 2016
12. Rocchesso, D., Seran, S., Behrendt, F., Bernardini, N., Bresin, R., Eckel, G., Frantinovic, K., Hermann, T., Pauletto, S., Susini, P., Visell, Y.: Sonic interaction design: sound, information and experience. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems, pp. 3969–3972 (2008)
13. Rocha, F.: Questões de Performance em Obras Eletrônicas Mistas. XX Congresso da ANPPOM, Goiânia. ANPPOM (2010)
14. Sharp, H., Rogers, Y., Preece, J.: Design de Interação: além da interação homemcomputador. 3a edição. ed. [S.l.], Bookman (2013)
15. TUIO. <http://www.tuio.org>. Accessed 09 Oct 2016