Banging Interaction in Ubiquitous Music

Damián Keller[1,2,4(✉)] , Azeema Yaseen[3] , Marcello Messina[1,2,4,5,6] ,
Sutirtha Chakraborty[3] , Luzilei Aliel[4] , Victor Lazzarini[3,4] ,
Ivan Simurra[1,2] , and Joseph Timoney[3,4]

[1] Amazon Centre for Music Research (NAP), Federal University of Acre,
Rio Branco, Brazil
`marcello@ccta.ufpb.br, ivan.simurra@ufac.br`
[2] Federal University of Paraíba, João Pessoa, Brazil
`dkeller@ccrma.stanford.edu`
[3] Audio and Music Computing Group, Maynooth University, Maynooth, Ireland
`{azeema.yaseen.2020,sutirtha.chakraborty.2019}@mumail.ie,`
`{victor.lazzarini,Joseph.Timoney}@mu.ie`
[4] Ubiquitous Music Group, Rio Branco, Acre, Brazil
[5] Southern Federal University (SFEDU), Rostov-on-Don, Russia
[6] Federal University of Acre, Rio Branco, Brazil

Abstract. We explore a new conceptual framework for interaction design in the context of ubiquitous musical activities, *banging interaction*. Our proposal entails a shift of focus from instruments, tools or device-oriented approaches to a timbre-led design practice that targets the potential relationships between sonic and multimodal qualities and the local available technological resources. One thread of this perspective envisages more flexible usage through the application of adaptive techniques. Another thread engages with recent advances in mid-air interaction, hinting at multimodal sensing. Yet another thread involves the incorporation of an aesthetically pliable approach to sonic materials, materialised through timbre as a target, rather than 'notes', 'instruments' or 'orchestras'. We present two examples of deployments: The prototype *mixDroid*, one of the first systems to employ mobile devices in everyday settings for creative purposes, and the prototype *Dynamic Drum Collective*, an adaptive implementation for percussive sources based on visual tokens and whole-body interaction.

Keywords: Ubiquitous music · ubimus · Adaptive interaction · Mid-air interaction · Timbre-based design · Time tagging · Dynamic Drum Collective

1 Introduction

Most of what has been done in musical interaction design during the last forty years has been tinged by a few basic assumptions: 1. Musical interaction targets professional musicians; 2. Music is made in professional artistic venues; 3. Musical activities are synchronous; and 4. Musical interaction is an extension of a

universal acoustic-instrumental way of thinking.[1] Until the turn of the century, some of these assumptions may have been enforced through a concentration of technical resources in a few major research and corporate facilities located in the central countries. But the expansion of the access to consumer-level technology in peripheral countries and particularly the emergence communities of practice that foster open-source support and production of know-how have triggered a radical change in the playing field while pushing for a democratisation of the access to creative music making. This is the context that enabled the emergence of the Ubiquitous Music community (ubimus) [12].

Emerging from the ongoing discussions on the theoretical underpinnings of ubimus design, current ubimus efforts feature a healthy balance between the speculative efforts and the emerging musical practices [11,16,22], the hands-on approaches to implementation and deployment of technological support   [17] and the artistic explorations that target fast-changing and firmly situated social needs [7,20,21]. Two ubimus strategies that have gathered traction entail the usage of transitional settings for casual interactions and an enhanced support for the generation and selection of timbre-based sonic resources. Motivated by this emerging trend, this paper explores the intersection of timbre-oriented design thinking with the harsh requirements of casual interaction in transitional settings. We summarise and discuss a body of experimental work developed within the first wave of ubimus projects that encompasses iterated cycles of design, deployment and assessment of several prototypes based on the creative-action metaphor Time Tagging, a design strategy tailored for audio-mixing in the field. We highlight the caveats and contributions of this approach in light of recent similar practices deployed in everyday settings. We also present a new creative-action metaphor that targets whole-body interaction, the Dynamic Drum Collective.

This family of prototypes features the ability to adapt to the particularities of the stakeholders' movements. In a sense, the software is tuned by using it. The automated adjustments are kept aligned with the specific characteristics and the spatial relationships afforded by the participant's body. These two metaphors, Time Tagging and the Dynamic Drum Collective, serve as examples of a new perspective on ubimus design, Banging Interaction. We consider three components of this framework, propose a conceptual representation, discuss the methodological implications and complete our argumentation with pointers to future applications.

2   Components of Banging Interaction

This section of the paper engages with three components of the banging-interaction conceptual framework: adaptive interaction, mid-air interaction and

---

[1] By "acoustic-instrumental thinking" we describe the social and technological approaches attached to the European instrumental musical tradition of the nineteenth century (i.e. the separation between the audience, the composer and the performer; the adoption of domain-specific knowledge as a requirement for music making; the enforcement of the myth of the genius as the source of creativity embodied by the virtuoso player, conductor or composer).

timbre-led design. Each thread is motivated by different factors that were manifest in previous ubimus projects but remained impervious to a unified approach. Adaptive design entails the usage of computational techniques to adjust the response of the tools to the behavioural and ergonomic characteristics of the subjects. An adaptive tool is the behavioural opposite of the instrument. An instrument, depending on the complexity of its functionality, frequently demands a considerable reshaping of human attitude and behaviour for its usage. Simple instruments, like the screwdriver, require little amount of training or bookkeeping. Complex instruments, like the acoustic string ensemble, involve a large amount of documentation and specialised knowledge to yield usable results.[2] Contrastingly, adaptive tools change their own behaviours as a result of a history of human-tool interactions [10]. Thus, their computational components tend to adapt to the human needs instead of forcing the human stakeholder into movements or cognitive patterns established by the design of the tool (as it is usually the case in instrumentally oriented design). Current ubimus ecosystems may tend toward a balance between the application of adaptive strategies and the incorporation of extant knowledge and functionalities. This approach is needed to enable the support of established musical practices which are at the core of most musical training, in tandem with the encouragement of new ways of music making. These aspects of musical interaction are, of course, addressed by the ongoing ubimus efforts and reflections. The last section of the paper will propose future developments along these lines.

Mid-air interaction [4] gained importance within the context of the restrictions imposed by the world pandemic. Contactless interfaces are particularly welcome in transitional settings involving unscreened and casual participations [15]. Despite the flexibility of touchless interaction, various challenges involving the drastic reduction of tactile feedback still need to be addressed [6]. Ultrasonic haptic technologies furnish a path to partially solve these issues, but they are still work-in-progress [24]. Within the context of Case 2 we will deal with these caveats and point to avenues for the future development of haptic-enabled banging-interaction endeavours.

The third component of banging interaction is timbre-led design. A recurring issue in ubimus deployments entails enlarging the metaphors' creative potential without increasing the cognitive effort involved in their usage. This issue is particularly problematic in transitional settings. When faced with open-ended activities and resources, some stakeholders tend to apply preconceived notions of sound making [18]. A typical reaction to flexible temporal organisations is "I want a beat, how do I make a beat?"; when dealing with environmental sounds, subjects may ask "Where is the violin?"; and their most usual general attitude toward creative tasks is "I do not know music", *music* meaning symbols writ-

---

[2] The temporal investment to achieve expertise has been reported at approximately ten thousand hours. The investment required to apply pre-existing knowledge to a different domain is still an open question. In any case, when dealing with casual interaction to require extensive musical training means to exclude a large number of potential participants.

ten on a staff. Thus, there is a gap between the target of the creative-action metaphors, i.e., enhancing the creative potential of the participants, and the stereotypes that some participants bring to the experimental sessions. Interestingly, according to the findings reported by [18], this form of cognitive bias is less pronounced in subjects that lack musical training than in those who went through a conservatory-oriented instrumental instruction.

To avoid the problems faced by an open-access policy to resources, a possible alternative may involve the reduction of the type of sonic materials employed by the creative-action metaphor. During the last century, some composers incorporated restrictions according to their personal style. A translation of this strategy into computational terms has triggered efforts in style emulations and genre classifications [8]. This approach is well suited for musicological investigations of the extant music repertoire. But it may be problematic when applied to creative endeavours. What aspects of a personal style are applicable across genres and beyond the needs of an isolated practitioner? How could this form of genre-oriented design foster musical innovation? And more generally, why adopt a design aligned with an established musical genre if a key objective of ubimus frameworks entails fostering the exploration of new forms of music making?[3]

A possible solution to this conundrum is furnished by timbre-oriented ubimus design. Timbre-driven design avoids the logic of acoustic-instrumental thinking. The standard musical interaction approach asks "What can I do with this instrument?" A very likely outcome is *functional fixedness.*[4] A colloquial way of describing this situation is illustrated by the saying "If all you have is a hammer, everything looks like a nail." Rather than starting the design process from a preconceived musical infrastructure, timbre-based design entails considering the characteristics of the targeted sonic resources. The required knowledge can be collected through various techniques, including human-based, machine-based or hybrid sonic analysis involving annotation procedures. The resource-selection stage entails the adoption of criteria aligned with the potential contexts of deployment. For instance, impact sonic materials tend to be amenable to isolated actions; Sustained sonic textures may afford continuous parametric-control strategies; And periodic temporal organisations may work well with iterated-control schemes. These strategies could eventually be layered and organised as composites that incorporate flexible procedural patterns.

The next two sections feature two examples of creative-action metaphors projected for everyday settings. Both support percussive sources. They can drive either audio-synthesis or processing techniques. They can be used for individual and/or collective activities. They are transparent to aesthetic choices and do not demand domain-specific knowledge or intensive training. Each case features a succinct description of the prototypes, a schematic report of the procedures and materials and a summary of results. Despite their positive assessment by

---

[3] There is an ongoing discussion of these issues within the ubimus community, with partial but promising results. See [16] for various examples.

[4] "A tendency to perceive an object only in terms of its common use." https://dictionary.apa.org/functional-fixedness..

non-musicians in the context of casual interactive usage, there are some limitations that need to be considered. The last section of the paper addresses their implications for design and suggests methodological strategies to overcome part of the caveats.

3   Case 1: Time Tagging

The time-tagging creative-action metaphor has been deployed in various contexts, fostering the investigation of the potential of ubimus technologies to boost engagement and participation in everyday musical activities. Two generations of prototypes were designed and deployed. The mixDroid 1.0 prototype lets users combine sounds by means of a virtual keyboard with nine touch-sensitive buttons on the mobile-device touchscreen. The mixing activity is based on triggering sounds through buttons while keeping track of trigger times. Since the control is limited to a single parameter, namely, time, the skills and training required are way less specific than those required to use an acoustic instrument. Furthermore, they do not depend on a symbolic system to be learned (e.g., common-practice music notation) and can be based on the characteristics of the sound material employed. Such a mechanism allows for the fast playback of up to nine sounds, depending exclusively on the sound matrix built during the selection activity, which lets the user assign each sample to each button of the interface.

As an initial validation process, the authors used an emulation of the first-generation mixDroid prototype (mixDroid 1G) for the creation of a complete musical work. The procedure encompassed several mixing sessions. The mixDroid 1G prototype was used in the emulation mode on a laptop computer and was activated through pointing and clicking with an optical mouse. Several dozens sound samples were used, with durations ranging from less than a second to approximately two minutes. The temporal structure of the mix followed the temporal characteristics of the sonic materials (biophonic sounds). The result was a seven-minute stereo track, Green Canopy On The Road, that we regard as the first documented ubiquitous music work.

Focusing on the demands of "lay participants"[5] in everyday contexts, a second study [3] consisted of creative activities in public (a shopping mall, a busy street and a quiet area featuring biophonic sounds) and private settings (the homes of each participant and a studio facility). Six subjects participated in 47 mixing sessions using samples collected at two outdoor sites comprising urban sounds and biophonic sources. Creativity support was evaluated by means of a creative-experience protocol encompassing six factors: productivity, expressiveness, explorability, enjoyment, concentration, and collaboration. Outdoor ses-

---

[5] Our interest for such a composite and poorly described social group is rooted in Milton Babbitt's (1958) well-known essay 'Who Cares if You Listen?' [23], and needs to be understood as a critique of the overt elitism incorporated in some narratives of musical creative practice. We decided to keep using the adjective 'lay' with reference to Babbitt's exhaustive mention of the 'layman', despite the fact that we find the term highly charged and stereotyped.

sions yielded higher scores in productivity, explorability, concentration and collaboration when compared to studio sessions. Compound effects of sound sample type and activity location were observed in the explorability factor when biophonic sound samples were used. Similar effects were detected on explorability, productivity and concentration in the conditions employing urban sounds.

A third study made use of recorded vocal samples created by the participants. In order to optimise the effects of place and activity type, three conditions were studied: in terms of place, domestic or commercial settings; in terms of activity type, imitative mixes or original creations; additionally, body posture, i.e., composing the mix while either standing or sitting, was also considered. Ten subjects took part in an experiment encompassing 40 interaction sessions using mixDroid. Subjects created mixes and assessed their experiences through a modified version of the CSI protocol applied in the previous studies. Explorability and collaboration factors yielded superior scores when the activities were carried out in domestic settings.

The results highlighted the impact of the venue on the support of everyday creative experiences. Outdoor spaces were preferred by the participants of the second study, while domestic settings got slightly higher ratings in the third study. While the profile of the subjects impacted the outcome of the third study, this trend was not confirmed by the results of the second study. Hence, the main conclusion to be drawn from these studies points to the impact of the venue on the subjects' evaluation of the creative experience. Both their ability to explore the potential of the support metaphor and their ability to collaborate were boosted by domestic settings and by outdoor settings.

The second generation of mixDroid prototypes features a new interaction mechanism: the stripe The stripe acts as a functional unit that features both interaction support and audio manipulation. This metaphor ties to the sonic sample the functionality previously linked to the audio channel in analogue systems. The objective is to allow for synchronous interaction with a large number of elements to overcome the screen-size limitations of small devices. Stripes enable mixing using both hands. The amount of active stripes depends on the device's computing power and on the participant's cognitive abilities. Thus, similarly to previous time-tagging implementations, devices with low computational resources can be used for complex creative activities in everyday contexts.

The stripe acts as an entry point to the sound data. Each stripe displays basic information on the sonic sample being handled, including the file name, the total running time, the current time and the execution state Each sound file linked to a stripe is processed independently. By linking the interaction mechanism with the sound sample, the stripe releases the user from the requirement of dealing with multiple samples as a block (as it is the case in the mixing-console metaphor that has the audio channel as its basic functional unit). Synchronous mixing of multiple sound sources is supported without compromising the parametric independence of each source. From the perspective of the user, sounds that demand fast interaction can be placed on stripes that are close to each other.

This flexibility, combined with the ability to select stripes though scrolling, is likely to grant quick access to a large number of sonic items.

The time-tagging metaphor can be subsumed under the general banging interaction framework, in that (1) it constitutes an example of adaptive design, tailored to fit the specificities and characteristics of an user base comprising non-musicians; (2) it replaces the metric orthodoxies of the acoustic-instrumental paradigm with a timbre-led type of thinking, where both the characteristics of the sonic material and the environmental local cues influence the temporalities of the musical outcomes; (3) in terms of mid-air interaction, the existing mixDroid prototypes are based on current touchscreen functionalities and, therefore, on contact-based tactile interaction, however, contactless mechanisms (using existing mobile phone functionalities such as hand-proximity and device-position sensors) might offer an interesting pathway for future mixDroid implementations.

4 Case 2: Dynamic Drum Collective

Whole-body touchless metaphors for music making encompass metaphors for executive tasks by direct control and adaptive metaphors projected for assistance. The first class may involve performative or recreational usage while the second type may support learning through cycles of parametric adjustments and user feedback. A direct-manipulation approach is usually applied for open-ended activities. This section provides a description of a new prototype designed and developed within the framework of banging interaction, the Dynamic Drum Collective.

The Dynamic Drum Collective is a camera-based gesture recognition system that enables individual and collaborative control of events aiming at both co-located and distributed collective music making with percussive sounds. The system detects the presence of the human body, tracks gestures and maps them to parameters of audio synthesis and processing engines. To promote collaboration, a client-server architecture was implemented to connect users remotely on multiple machines. The prototype akin to a virtual and multitimbral drum demands a fairly low usage of bandwidth.

4.1 Adaptive Visual Tracking

To enable a consistent and stable visual guide for their actions, the users may configure a fixed arrangement of colour-coded adaptive visual tokens. These referents act as anchors or opportunities for action and are dynamically aligned to the spatial displacements of the subject [2].

Individual-usage mode involves a combination of adaptive visual tokens and sonic feedback through the detection of limb movements by means of a pose-estimation algorithm. Input comes from a standard consumer-level video camera (in the case of this study a built-in model bundled on a laptop computer). The video images are analysed using a pose-estimation algorithm which drives the visual and audio feedback provided to the user.

Calibration is flexible. The subjects set up the visual referents using both hands. To handle various camera positions, body characteristics and camera-player distances, the referents are dynamically adjusted.

During displacements, the tracking system adjusts the visualisation to the body position to compensate for changes of angle and distance from the camera. Regardless of the relative subject-camera distance, playing remains compatible with dance-like motions and gesture mimicry. Thus participants can move freely without losing their sonic-spatial frame of reference. Assigned colours complement the information provided by the sonic outcome, indicating if a movement has been detected and which visual token is active. Consequently, each occurrence of an event is reinforced by an appropriate visual cue.

As a downside, while simple and clear visual feedback may scaffold the sonic results it requires using displays that are large and easily visible from a distance. The visual tokens move along the X-Y coordinates as the user changes positions. A deep-learning model [19] is used for the synchronous detection of the three-dimensional body positions. Displacements are represented by coordinate changes of the spatial referents for each video frame The coordinates of the hips are taken as anchors to track the centre of the body (orange marker). The distance between the subject's ears is used to estimate the subject's head movements toward and away from the camera (red marker). This cue triggers a new calculation of body distance, allowing the system to keep the visualisation within scale.

A queue-replacement technique was implemented to track the movements of each arm. There are two dedicated buffers: 1. The 'speed buffer' is updated synchronously to reflect the spatial displacements of the hands. Two consecutive frames are used to estimate the speed of motion. Index-finger movements are denoted as event triggers (purple marker) and are used to track grasping, pinching and striking actions. When an index finger 'touches' a visual token at a speed exceeding a threshold value, the event is processed as a 'hit'. Depending where the hit occurs, a MIDI message is triggered. The event's intensity (MIDI velocity) is calculated from the mean value of the speed buffer. Thus, if a limb position changes rapidly the values in the speed buffer increase accordingly, resulting in a louder event.

4.2   Sonic Rendering

In this implementation of the Dynamic Drum Collective, the percussive sounds are rendered locally and the remote usage relies on exchanges of MIDI control data through a dedicated server architecture. The first iterations of prototypes use loopMIDI to create MIDI-ports to forward the data to a DAW (in this case MixCraft) that renders the sound.

4.3   Deployments

First Prototype. The first prototype of Dynamic Drum Collective was presented live at the Ircam Workshop 2022. [5]. The highlights were: i) A calibration

mode featuring the action of grasping or pinching to handle the visual referents of a drum kit, ii) Support for adaptive movement tracking to enable consistent sonic feedback during whole-body interaction, and iii) Multi-user collaboration through network connectivity. During the live demonstration, the system was used by two volunteers from the audience including one professional drummer and a non-musician. While the system performed fairly well with respect to latency, there were some issues caused by the characteristics of the venue. Varying light conditions caused visual jitter. The informational noise added extra time to the image-processing cycle. The accumulated delay was larger than during the earlier laboratory trials, compromising the temporal alignment of the sonic feedback. This caveat did not preclude the interactions among the participants but it points to the increased difficulties posed by the everyday-settings deployments.

Second Prototype. This version of Dynamic Drum Collective uses colours as visual tokens to avoid the display of images of acoustic instruments. The system was used by musicians who were also experienced percussionists. During the exploration session, our aim was to understand how well the user dealt with latency or missed targets during playing. This session helped us to evaluate the relationship between the GUI and the body movements to identify major flaws and potentially useful features.

Deployment. Testing was divided into two stages: (i) free exploration and (ii) semi-structured tasks. A professional musician with ample experience and training in technological design participated in the testing sessions using the system. During the first stage, the participant did not receive any information on how to use the tool. The objective was to focus on the limitations of the support for casual interaction by tracking the time needed to reach a basic understanding of its usage. Subsequently, we observed whether he was comfortable doing targeted activities with the drum prototype. This stage allowed us to gather information on the specific design dimensions and demands of an unfamiliar ubimus ecosystem, providing us a notion of some of the major hurdles to be encountered in field deployments.

During the initial exploration phase, the participants were given no visual feedback of their movements (no-visuals condition). They could only see the bounded regions for triggering sound. The instruction was: "a 'hit event' occurs by moving the index finger inside the visual tokens". The lack of spatial referents made it hard for the subjects to align the visual tokens in relation to their own position. According to them, this condition had a negative impact on the accuracy of their actions. Our analysis of the footage corroborated these observations.

Participants used the square 'Calibration box' to position the adaptive visual tokens. One hand is employed to select a kit and the other arm serves to drag the token position to the desired location. The top corners of the token change from blue to cyan when activated. Session 2. The subsequent exploration session

involved playing short rhythmic phrases (of approximately 5 to 10 s) until completing ten iterations for each session. When a percussive event is detected, the background changes to the colour of the token just touched.

Results. Once the visual feedback was enabled, the body keypoints were tracked in real time (with-visuals condition). According to the subjects, the cross-modal feedback helped them to keep track of the relationship between their movements and the motion targets. They also mentioned that the markers overlaid on the hands improved the quality of the experience.

Some participants found it hard to maintain steady sequences at a fast tempo. Large movements (such as those applied for slow rhythms) seem to work better than small gestures. The smaller movements of the fast sequences sometimes cause drop-outs due to system latency or to pose-estimation inaccuracies. The participant also mentioned that sidewise hand movements were better tracked by the system.

We observed that the execution of the rhythmic phrases was easily achievable when the tempo ranged between 60 to 125 BPM (for the simple sequences) or 60 to 90 BPM (for the complex ones). Most subjects could maintain the tempo in the with-visuals condition (90%). Without visual feedback, the subjects missed 3 trials out of 10 (70%). Errors included missed targets and double hits.

An open-ended, exploratory session was proposed as a way to encourage alternative interaction strategies. This stage allowed the participants to fine-tune their gestural approach while exploring both very large and very small motions. One of the activities involved playing while walking. The subjects were comfortable with the demands of this activity.

Third Prototype. The most recent prototype of this project expands the previous functionalities with four visual tokens to interact with percussive sounds, provisionally featuring a hihat, a cymbal, a snare drum and a kick drum. The gestural space remains the same. The focus of this study was to explore the potential of mid-air interaction for casual musical activities with a larger pool of participants.

To reduce the processing load and to enable consistent movement of visual tokens along with human body movement, every time a change in reference point is detected it is compared with a threshold value. If the tracked value is bigger than the threshold, coordinates of visual tokens are recalculated otherwise no change is made. So, the new prototype is more consistent and only process calculated coordinates when a user's step size is a large scale movement (Fig. 1).

Two laptops were used to conduct the study of prototype V3. Both machines were running 64-bit Windows 10 with an AMD Ryzen and an Intel Core i7 CPU, 16 GB and 32 GB of installed RAM, respectively. The video streams were captured with 0.3 and 0.9 MP camera sensors with a resolution of 640×480 pixels at 25 frames per second (fps) and 1280×720 pixels at 30 fps.

Deployment. This study was carried out with 15 non-musicians. Except for one person, all the participants were experienced in using computers and information
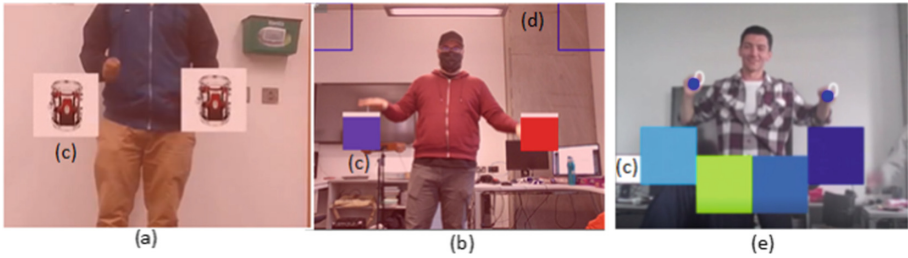
Fig. 1. Three prototypes of the Dynamic Drum Collective metaphor designed for individual and group musical interactions, featuring visual tokens for percussive sounds and boxes for calibration.

technology (ranging from 3 to 10 years of experience). Seven participants had no formal music training. Of those seven, two were not formally trained but had previous informal experiences in music making. The rest had less than two years of formal music education, including self-instruction. None of them described themselves as active musicians.

Each session was divided in two parts: i) individual interaction and ii) collaborative interaction. At the beginning of the session, every individual spent some time exploring the prototype freely. Once they were comfortable with the basic functionalities, they were asked to attempt to synchronise their actions with the other participants. Beyond that, they were not given any specific instructions on the types of rhythmic sequences to be played.

Results. We observed a learning pattern while the participants were trying to synchronise. As they started the activity, they followed each other triggering one percussive sound at a time. When they felt satisfied with a simple sequence, they introduced other events to their patterns. During the collaborative task, the complexity and alignment of their performance was correlated to the amount of time invested in the activity.

5   Discussion of Results and Future Ubimus Deployments

Banging interaction targets the development of creative-action metaphors with an emphasis on the usage of percussive sources and audio-processing tools tailored for everyday settings. These settings demand flexible configurations to respond to fast-changing conditions and highly varied participant profiles. Three components of banging interaction - adaptive interaction, mid-air techniques and timbre-led design - complement the standard methods of ubimus design. Our proposal encompasses repurposed mobile consumer devices, embedded and networked equipment, without excluding the deployment of robotic tools [17]. This framework also addresses aspects of design that have resisted an integrated conceptual and methodological treatment, highlighting: unscreened and casual forms of music making, transitional noisy and fast-changing settings, reduced

tactile feedback and usage of the whole body as a potential source of musically meaningful information.

The two creative-action metaphors featured in this paper share characteristics that make them especially suited for banging-interaction design. They support sonic sources with temporal profiles that can be roughly classified as impact sounds. They incorporate either audio-synthesis or audio-processing techniques that do not demand a high computational cost, making them amenable to usage through network infrastructure. Locally, they support synchronous usage. They may also be employed in the context of individual or collective musical activities. Given their reliance on local resources, they tend to be scalable and may be integrated into the Internet of Musical Things or the Internet of Musical Stuff, expanding the extant IoMuSt initiatives [22]. As long as the musical activities make use of impact sounds, they remain fully transparent to various aesthetic perspectives. Last but not least, they do not demand domain-specific knowledge or specialised intensive training.

One of the components of banging interaction entails a change of perspective regarding the selection of sonic resources. Rather than starting the design process from a preconceived musical infrastructure, typically represented by the figure of the "musical instrument", timbre-based design engages with the characteristics of the targeted relational properties, i.e. the possibilities afforded by the multimodal aspects of the musical activity [13]. The knowledge required by the resource-selection process can be obtained through various techniques, including human-based heuristics, machine-based decision-making processes involving adaptive techniques and hybrid forms of sonic analysis and annotation that may combine simple and more elaborate interaction-design patterns [9]. As hinted by the results of the deployed metaphors, resource selection should be aligned with the potential contexts of usage. Thus, impact sonic materials are well suited to isolated gestures but other forms of sonic organisation may not be amenable to this family of metaphors.

A potential limitation of the current consumer-grade technologies available for banging interaction is the reduction of tactile information and feedback. Standard touchscreens do not feature pressure sensing or proximity sensing. Consumer-level video cameras are useful tools for capturing static and moving elements of visual scenes but haptic aspects remain elusive and haptic feedback demands other types of technologies. Nevertheless, the landscape of technical resources for everyday contexts is changing quickly. Touch-based interaction can be potentially expanded by the incorporation of pseudo-haptics. Through the use of proximity sensors and more refined techniques of contact-based tracking, it may be possible to repurpose the consumer-level mobile devices to act as parts of haptic-enabled ubimus ecosystems rather than as isolated objects or as emulations of acoustic instruments. Contact-based interaction involving personal wearable devices may still have its place within the panoply of everyday creativity-oriented technologies. Furthermore, touchless whole-body interaction may soon incorporate ultrasonic haptic feedback. This functionality is not yet featured in consumer-level portable devices but given the rate of expansion of

the current computational platforms, it would not be surprising to find affordable equipment to test and deploy these techniques within the context of the DIY initiatives.

A final aspect to consider, that has gained centrality within the context of post-2020 musical practices, is the potential ecological footprint of the incorporation of technology in everyday settings (either in transitional settings with large numbers of casual users or in domestic settings with access restricted to the home-dwellers but with potentially negative consequences on privacy and well-being). Whole-body, contactless interaction furnishes a viable solution to the health risks involved in casual participation by unscreened stakeholders in transitional settings. The beauty of the Dynamic Drum Collective metaphor is that it does not demand anything not already available in shopping malls, bus stations, airports or leisure spaces: consumer-level computers and standard built-in video-cameras. Sound may be delivered to personal headphones or it can also be rendered through standard loudspeakers. Given a MIDI-based audio synthesis tool, any mobile or embedded hardware can be employed. Thus, the constraints of its deployment are more related to the social implications of music making in public spaces than to the lack of access to hardware. Social interaction in public settings is a highly complex and delicate issue. Hence, the implications of deploying technology for music making need to be carefully contextualised regarding the cultural expectations around the use of transitional settings. Future banging-interaction studies will have to address this issue.

Furthermore, domestic ubimus is emerging as a viable research area with its own methods and potential problems [14]. Banging interaction may be easily inserted as a thread of domestic-ubimus endeavours, but a few words of caution are required regarding the delivery of sound and the usage of camera-based systems at home. The location of the system needs to be carefully chosen to foster engagement while avoiding the disruption of daily routines. This demands ethnological studies in diverse cultural contexts to establish not only the types of usage that are out of bonds but also to determine potential conflicts with other daily activities. We have already seen too many examples of mindless computing [1]. An objective of ubimus is the expansion of wellbeing. This involves a responsible deployment of technological resources which can be used to enable mutual engagements, to boost positive attitudes and to complement and encourage healthy habits. Despite all these potentially beneficial aspects, the same technological resources could eventually feed the unlimited voracity of corporate greed. A conscious and alert posture regarding the preservation of intimate and surveillance-free spaces is one of the pressing needs of post-2020 ubiquitous computing.

References

1. Adams, A.T., Costa, J., Jung, M.F., Choudhury, T.: Mindless computing: designing technologies to subtly influence behavior. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2015, pp. 719–730. Association for Computing Machinery, New York (2015). https://doi.org/10.1145/2750858.2805843

2. Algethami, N., Redfern, S.: A robust tracking-by-detection algorithm using adaptive accumulated frame differencing and corner features. J. Imaging **6**(4), 6040025. https://doi.org/10.3390/jimaging6040025, https://www.mdpi.com/2313-433X/6/4/25

3. Babbitt, M.: Who cares if you listen? (1958). In: Strunk, O., Treitler, L., Morgan, R.P. (eds.) Source Readings in Music History, vol. 7: The Twentieth Century, pp. 35–41. Norton, New York (1998). http://www.palestrant.com/babbitt.html

4. Bolt, R.A.: "put-that-there": Voice and gesture at the graphics interface. In: Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques, pp. 262–270. ACM, New York (1980)

5. Chakraborty, S., Yaseen, A., Timoney, J., Lazzarini, V., Keller, D.: Adaptive touchless whole-body interaction for casual ubiquitous musical activities. In: Proceedings of the International Computer Music Conference (ICMC2022), pp. 132–138. University of Limerick, Limerick (2022)

6. Chattopadhyay, D.: Understanding interaction mechanics in touchless target selection. PhD Thesis in Computer Science. Indiana University, Bloomington, IN (2016)

7. Connors, T.: From the edge. In: Proceedings of the 11th Workshop on Ubiquitous Music (UbiMus 2021), pp. 17–19. g-ubimus (2021)

8. Cope, D.: Virtual Music: Computer Synthesis of Musical Style. MIT Press, Cambridge (2004). http://books.google.com.br/books?id=dJ7Numf8fwMC

9. Flores, L.V., Pimenta, M.S., Keller, D.: Patterns of musical interaction with computing devices. Cadernos de Informática **8**(2), 68–81 (2014). http://compmus.ime.usp.br/ubimus/pt-br/node/33

10. Graham-Knight, K., Tzanetakis, G.: Adaptive music technology: history and future perspectives. In: Proceedings of Internationl Computer Music Conference (ICMC 2015), pp. 416–419. CEMI, University of North Texas, Denton (2015)

11. Keller, D., Aliel, L., Filho, M.C., Costalonga, L.: Toward ubimus philosophical frameworks. Open Phil. **4**(1), 353–371 (2021). https://doi.org/10.1515/opphil-2020-0188

12. Keller, D., Lazzarini, V., Pimenta, M.S.: Ubiquitous Music, vol. XXVIII. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-319-11152-0

13. Keller, D., Otero, N., Lazzarini, V., Pimenta, M.S., Lima, M.H., Johann, M., Costalonga, L.: Relational properties in interaction aesthetics: The ubiquitous music turn. In: Ng, K., Bowen, J.P., McDaid, S. (eds.) Proceedings of the Electronic Visualisation and the Arts Conference (EVA 2014), , London, UK, 8–14 July 2014. BCS, Computer Arts Society Specialist Group, London. https://doi.org/10.1426/ewic/eva2104.25

14. Keller, D., Simurra, I., Messina, M., Neiva, T.M., Tedesco, S., Mesz, B.: Domestic ubimus. EAI Endorsed Trans. Creative Technol. **9**(30), 173493 (2022). https://doi.org/10.4108/eai.22-2-2022.173493

15. Keller, D., Gomes, C., Aliel, L.: The handy metaphor: bimanual, touchless interaction for the internet of musical things. J. New Music Res. **48**(4), 385–396 (2019). https://doi.org/10.1080/09298215.2019.1654525

16. Kramann, G.: Composing by laypeople: a broader perspective provided by arithmetic operation grammar. Comput. Music J. **44**(1), 17–34 (2020)
17. Lazzarini, V., Keller, D., Otero, N., Turchet, L.: Ubiquitous Music Ecologies. Taylor & Francis (Routledge), London (2020)
18. Lima, M.H., Keller, D., Pimenta, M.S., Lazzarini, V., Miletto, E.M.: Creativity-centred design for ubiquitous musical activities: two case studies. J. Music Technol. Educ. 5(2), 195–222 (2012). https://doi.org/10.1386/jmte.5.2.195_1
19. Lugaresi, C., et al.: Mediapipe: a framework for building perception pipelines. arXiv preprint arXiv:1906.08172 (2019)
20. Mackay, R., Rodríguez, J., Rodríguez, R., Blink, D., López, P.J.: Networking the flight of the monarchs (performance). In: Proceedings of the Ubiquitous Music Symposium (UbiMus 2022), pp. 141–142. g-ubimus, Curitiba, PR (2022)
21. Messina, M., Aliel, L.: Ubiquitous music, gelassenheit and the metaphysics of presence: Hijacking the live score piece ntrallazzu 4. In: Proceedings of the Workshop on Ubiquitous Music (UbiMus2019), Marseille, France, 14–18 October 2019, pp. 685–695. Proceedings of the 14th International Symposium on CMMR, Marseille (2019)
22. Messina, M., Célio Filho, M., Mejía, C.M.G., Keller, D., Aliel, L., Simurra, I.: A internet do bagulho musical (internet of musical stuff)-iomust. In: Proceedings of the Ubiquitous Music Symposium (ubimus2022), pp. 85–93. g-ubimus (2022)
23. Pinheiro da Silva, F., Pimenta, M.S., Lazzarini, V., Keller, D.: Time tagging in its niche: engagement, explorability and creative attention (a marcação temporal no seu nicho: Engajamento, explorabilidade e atenção criativa). Cadernos de Informática **8**(2) (2014). http://seer.ufrgs.br/cadernosdeinformatica
24. Subramanian, S., Seah, S.A., Shinoda, H., Hoggan, E., Corenthy, L.: Mid-air haptics and displays: Systems for un-instrumented mid-air interactions. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI2015). ACM, New York (2015)