

# Chapter 4

## Mediated Interactions and Musical Expression—A Survey

Dennis Reidsma, Mustafa Radha and Anton Nijholt

### 4.1 Introduction

The dawn of the information and electronics age has had a significant impact on music. Digital music creation has become a popular alternative to playing classical instruments, and in its various forms has taken a place as full-fledged class of instrument in its own right. Research into technological or digital instruments for musical expression is a fascinating field which, among other things, tries to facilitate musicians and to improve the art of musical expression. Such instruments broaden the available forms of musical expression and provide new modes for expression, described by some as a reinvention of the musician's proprioceptive perception (Benthien 2002; Kerckhove 1993). They can make musical expression and musical collaboration more accessible to non-musicians and/or serve as educational tools. Technology can also eliminate the boundaries of space and time in musical collaboration or performance, or enhance it, by providing new channels of interaction between performers or between performer and audience. Furthermore, technology in itself can be a collaborating partner in the form of a creative agent, co-authoring, helping or teaching its user. Finally, Beilharz brings forward the human desire for post-humanism and cyborgism in musical expression as a goal in itself to explore mediating technologies (Beilharz 2011).

In this chapter we will survey music technology through various lenses, exploring the qualities of technological instruments as tools, media and agents and investigating the micro-coordination processes that occur in musical collaboration, with the long range goal of creating better technological artifacts for music expression.

---

D. Reidsma (✉) · M. Radha · A. Nijholt  
Human Media Interaction, University of Twente, PO Box 217,  
7500 AE, Enschede, The Netherlands  
e-mail: d.reidsma@utwente.nl

M. Radha  
e-mail: mustafa.radha@gmail.com

A. Nijholt  
e-mail: a.nijholt@utwente.nl

As a starting point, the next section discusses our theoretical starting point for looking at the relation between (technological) instruments and their users.

## 4.2 Users and Musical Instruments: Conceptual Framework

Musical instruments are designed with the purpose of engaging, intriguing and appealing to humans. There are various categories of user of musical instruments (most importantly, *musicians* and *non-musicians*) who differ in their needs and interactions with music instruments (Coletta et al. 2008; Akkersdijk 2012; Beaton 2010).

Understanding how humans deal with technology, including music technology, and how technological artifacts influence our world, are elements of an ever-growing body of research. Verplank (2011) suggests to use a *tool-media-agent* model to understand the impact of (music) technology. Verbeek, building upon Ihde's post-phenomenological work (Ihde 1986), proposes three manifestations of technology: as a tool for a human; as a medium between two humans; or as an agent in itself, part of the social network (2005). An important aspect in each of these manifestations of human music interaction is the underlying *collaboration process*. Tatar (Lee et al. 2012; Tatar 2012) urges exploration of micro-level, situated actions and their broader outcomes, a process she coins as *micro-coordination*. Studying micro-coordination in interaction settings gives the designer a better understanding of the collaborative phenomenon and, as a consequence, better tools to design technology for the domain. In the music making domain, insight into this topic will enable us to invent better (mediating) instruments by accounting for the necessary micro-coordination, and to implement realistic music-creating agents to work with humans that are able to understand—and participate in—the necessary processes of coordination. We apply these models to musical instruments as follows.

**Tools**—Music instruments as tools merely serve as a means for their user to more easily achieve a goal. In Sect. 3, we will investigate the design of these tools in the light of different modes of interaction and how they satisfy various user needs. How can technology be used to enhance and create new interfaces for musical expression? Which technological modes of interaction can be exploited for musical expression? How can these modes of interactions serve the needs of different types of users?

**Media**—Instruments as media provide a channel of communication with other humans. Instruments of this type are mediators in social (human to human) interaction. In Sect. 4, we will look at how communication related to the context of collaborative musical expression can be realized. How can instruments capture the micro-coordination involved in collaborative musical expression? What micro-coordinating communication occurs in collaborative musical expression? How can technology be a mediating channel for this micro-coordination? How can technology augment the taxonomy of micro-coordination, enabling new ways of coordination?

**Agents**—An agent is an artifact possessing agency: it has beliefs, desires and intentions (Georgeff 1998), is part of the social structure as an entity in itself, and thus introduces the field of human-agent or human-robot interaction. We will study the design prerequisites for co-creative musical agents in Sect. 5.

### 4.3 Interaction Modes for Musical Expression: The Instrument as Tool

In this section, the technological instrument for musical expression will be discussed as a tool, investigating how new interaction modes enable new forms of musical expression. We will discuss different novel technological interaction modalities and how these complement the musician's needs, and look at how technology can support *non-musicians* in achieving music making experiences.

#### 4.3.1 Technology for Multimodal Music Interaction

Technological innovations allow for interaction modalities that go far beyond the simple keys of a digital piano. This section discusses three of the most important new modalities: tactile interfaces, gestural control, and brain-computer interaction for music making.

##### Tactile Interfaces

Many technological music instruments rely on touch, tangibility and haptics as control interfaces. The success of the tactile interface can be contributed to several factors.

The first reason is the ease of use or intuitiveness. Tactile interfaces exploit the user's natural affinities with vision and touch. We are used to explore the world through our eyes and manipulate it with our hands. Making this very real form of interaction with the world into an interface through augmented reality gives us many possibilities as is shown in many diverse projects (Fikkert et al. 2009; Jorda et al. 2007; Levin 2006; Patten 2002; Poupyrev et al. 2000; Raffa 2011).

Haptic interfaces may also invite the user into new forms of expression. Bill Verplank (Verplank et al. 2002) has done research into the expressive effects of haptic force feedback as a means of informing the user of their action. The power of haptics is illustrated in his *Plank* device, a slider with natural force feedback that resembles normality (the further you push the slider, the more resistance it produces). Such sophisticated interaction methods can be a powerful enhancement to the currently popular and successful touch-based and tangible instruments.

Another benefit is the extendable size of such interfaces. Many inventions like the *FeelSound* (Fikkert et al. 2009) and *reacTable* (Jorda et al. 2007) are table-sized



**Fig. 4.1** The reacTable interface (Jorda et al. 2007)

tactile interfaces designed to enable multiple musicians to work together on the same piece.

Not only the usability, but also the utility of the instruments can vary widely. While most tactile music interfaces look the same, their purposes are often different because of the adaptable nature of tactile interfaces. In illustration, the idea of a table-sized touch screen with tangible objects on them has been reused for composition (Fikkert et al. 2009), spectrograph manipulation (Levin 2006), rhythm creation (Raffa 2011), mixing (Poupyrev et al. 2000; Jorda et al. 2007) and harmony creation (Jorda and Alonso 2006). This promises the possibility of an all-round tactile table-top installation for musical expression in many forms.

The last aspect which especially sets apart tactile interfaces from the old-fashioned MIDI knob/button controller for the laptop, is the appealing interface for the audience. Paine (2009) emphasizes the need for interfaces that give more insight into the creative effort of the musician. He states that the laptop musician creates a distance barrier between him and the audience since his actions are not apparent for the audience. Projects like the reacTable (Jorda et al. 2007) (Fig. 4.1) serve to make the performance of the electronic musician more appealing.

## Gesture Interfaces

As computer vision and the detection of movement advance, gesture interfaces are more and more explored as a new mode of musical interaction. Gestures can be *quantified* to control parameters of the music, but can also be *sonified* directly into sound.

These two extremes can be generalized into a certain dimension of control mechanism design, *directness of control*. With indirect control, the output is only modified by the input. With direct control, there is a direct mapping of input to output.

An example of indirect control is the *Mappe per Affetti Erranti* (Coletta et al. 2008). The system monitors the movements of dancers on a stage and quantifies them into parameters that are applied to a preprogrammed music piece. The goal of the project is to let dancers conduct the music instead of the other way around. Virtual orchestra conducting interfaces (Borchers et al. 2004) and virtual conductors for real orchestras (Reidsma et al. 2008) share a similar mechanism. A non-gesture example of indirect control is the *Wayfaring Swarms* installation (Choi and Bargar 2011), a tabletop installation with music-generating swarms with which the user can interact to some extent.

Direct control gesture interfaces, in which the gestures are directly sonified, are scarce. A well-known realization of gesture sonification is the Theremin. According to Ward et al. (Ward and O'Modhrain 2008), the mastery of such instrument lies in the execution of managed movements. Goïna and Polotti (2008) investigate how gestures can be sonified in general by identifying the elementary components of gestures, working towards the idea of melting the virtuosity of dance and music into a single art, which confirms the importance of managed movement.

## Brain-Computer Interfaces

One of the latest additions to the repertoire of multimodal interaction technologies for HCI is that of Brain-Computer Interfacing: brain signals from the user are registered using, e.g., an EEG device, and signal processing techniques are used to obtain relevant information about the user's mental state (Nijholt et al. 2008). In its most primitive form, such an interface simply serves as a trigger which is activated through a specific kind of brain activity rather than through pushing a button. This allows musical expression to become available for people with physical disabilities that prevent them from playing regular instruments (Miranda 2006; Chew and Caspary 2011; Le Groux et al. 2010).

BCI interfaces have also been used for *direct control* of musical expression. Examples of this are the early work "Music for Solo Performer" by Lucier (Teitelbaum 1976) and "Staalhemel" (De Boeck 2011)—an installation in which a person walks under steel plates hung from the ceiling, and their brain signals are translated to small hammers that tap the steel plates according to the amplitudes of their brain signals. BCI based musical expression is also possible in multi-user settings. Sanger et al. (2012) looked at how collaboratively making music causes the brains of the musicians to synchronize. MoodMixer (Leslie and Mullen 2011) offers an EEG-based collaborative sonification of brain signals. Interestingly, these and similar project blur the distinction between direct control and indirect control. Consciously manipulating your own brain activity is possible, one can see what the effects are, relax, concentrate, think of something sad, think of something happy, etcetera. At the same time, brain activity is always there, it cannot be suppressed, and it can be manipulated in a feedback loop from the outside by exposing the user of the

instrument to changes in the environment. Then, sonification is ‘controlled’ by what the user experiences, but the user experience is simultaneously strongly controlled by the output of the sonification.

### 4.3.2 *Music Interaction for Non-Musicians*

These new interface technologies not only serve musicians, but are also usable for non-musicians and people who are still learning to play music. Two keys to engaging non-musicians are (1) making it easy and intuitive to make music and (2) giving the non-musician a taste of musical achievement.

Table-top installations have already been discussed as intuitive and natural interaction methods for musical expression. They therefore can appeal more to non-musicians than conventional instruments. Efforts in tangible interfaces focusing on teaching music are the MIROR project (MIROR 2012) and Melody-Morph (Rosenbaum 2008). Both projects exploit the loose building blocks of the tangible interface to make a *configurable interface*. This eliminates the learning curve associated with figuring out how the interface responds to the user’s actions, since the user can himself design the triggers and responses.

When technological instruments are used to offer the experience of a high level of musical performance to non-musicians without them (yet) having the real skills for such performance, it is called *simulated musical achievement*. An example of simulation of musical achievement are the Guitar Hero games (Harmonix Music Systems 2005). Miller (2009) explains the success of these games in the fact that they bear in them a form of *schizophrenic performance*, where the player, while not playing the instrument, does have the feeling of actually making the produced music. This separation between input and output (thus: schizophrenia) can essentially be generalized to the indirect control mechanisms discussed earlier. The levels of difficulty in the game represent the spectrum of directness. This principle can be applied to other interfaces such as gesture interfaces, where the interface can gradually transform from gesture parametrization into real sonification. Simulated musical achievement may in this way be used to motivate people to learn the actual skills for music making.

## 4.4 **Mediated Communication in Musical Expression: The Instrument as Medium**

A significant part of the research area concerned with musical expression focuses on the *collaborative* aspects of music-making. Many studies on this topic aim towards instruments that enable interesting forms of collaboration and interaction with other humans. In this section, we will look at research done to enable mediated human-human interaction in musical collaboration.

Communication in musical expression			
Micro-coordination in collaboration		Performer-audience communication	
Synchronization	Differentiation	Performance facilitation	Interaction
playing style -loudness -pitch -rhythm meta-information (spatial dislocation) -linking sounds to their creators meta-information (temporal dislocation) -awareness of action -annotation -shared representation -mutual modifiability	leader/follower role establishment leading -pitch change -loudness change -rhythm change following -following up on pitch -following up on rhythm -following up on loudness	virtuosity affective expression dance speech	entrainment
		Used channels	
		visual channel- non-verbal behaviour verbal channel- verbal behaviour auditory channel- musical behaviour	

Fig. 4.2 Taxonomy of interactions in musical expression as found in literature survey

During musical collaboration, traditionally there are three channels present, namely the visual, the auditory and the verbal channels (Beaton et al. 2010). The visual channel deals with body movements, the auditory are actions performed with the instrument (e.g. a drumming break) and the verbal channel contains only utterances and words.

The actions performed over these channels can also fall into different categories based on their intentional goals, forming a taxonomy of actions over two dimensions: channels and goals. As an analogue, Bianchi-Berthouze (2013) has developed a classification of the movements performed in gaming, which consist of task-controlling, task-facilitating, affective, role-facilitating and social movements. We will establish such a taxonomy in this section. The complete taxonomy as explained in this section can be seen in Fig. 4.2.

We first divide actions into two contexts: collaboration between musicians (micro-coordination) and performance towards the audience. Micro-coordination between musicians can happen for three different collaborative goals: the first is to achieve mutual synchronization between the musicians (e.g. entrainment, mutual awareness of meta-information, countdown) and the second is to dynamically change the music in a coordinated manner by establishing roles (leading, following). Between the musician and the audience, there are “monologue” communications by the musician, part of the performance, and interactive communications where the musician and audience work together to make the event.

4.4.1 Collaborative Micro-Coordination

This section concerns the micro-coordination that happens in cooperative music-making. We discriminate between two types of micro-coordination in musical expression: the first has a goal of synchronizing the collaborators with each other,



while the second type has an opposite goal: the goal of differentiating the music to create something new and to ensure a dynamic flow.

## Synchronization

Synchronization is done throughout the whole collaboration: it is used to initiate the music, to keep the different musicians in harmony and rhythm throughout the collaboration and finally to finish the created musical piece in a synchronized manner. Synchronization can occur over all 3 channels of communication.

Before starting musical expression, musicians have to communicate which piece of music to play (usually verbal) and perform a countdown of some sorts before commencing. This countdown can happen either verbally by counting, auditory by, for example, 4 hits on a percussion instrument or visually, for example by the leader through an exaggerated body movement before starting.

During musical expression, synchronization is needed constantly to work together. It appears that not only hearing each other playing music, but also being able to see each other while doing so, are important for synchronization. For example, Beaton (Beaton et al. 2010) found out that musicians could collaborate better with humans than with robotic performers because robotic performers lack visual feedback. Akkersdijk (2012) showed that the placement of a screen between musicians prevented them from synchronizing as well as they would when they can see each other.

The robotic improvisational Marimba-playing agent Shimon (Hoffman and Weinberg 2010) was designed based on research into meaningful movements for robots (Hoffman and Breazeal 2006, 2008; Hoffman et al. 2008) in order to emulate a human musician. Unique to this project is the effort put in the design of movements. The robotic arms that play the Marimba make elaborate movements to express several properties of the music: the motion size naturally denotes the loudness of a keystroke; the location of the gesture naturally corresponds to the pitch of the keystroke; the movements are exaggerated for further visibility for the other musicians. In another project, Varni et al. (2008) showed that humans collaborating with each other try to entrain their body movements to the external jointly established rhythm, and that this becomes more difficult when the humans are not able to see each other.

## Synchronization During Spatial Dislocation

The ten-hand piano (Barbosa 2008), part of the Public Sound Objects (PSO) project (Barbosa and Kaltenbrunner 2002), is envisioned as a piano playable by multiple people who do not necessarily see each other, in public spaces. The user interface represents different users as circles on the screen. The color of these circles denotes the pitch and the size denotes the loudness of the music produced by that person. This way, users who cannot see each other still can differentiate the mix of produced sounds in terms of their origin. In other words: this provides necessary



meta-information about the joint musical product, which is necessary for mutual collaboration. This helps synchronization as users are able to relate sounds to the persons producing them.

### Synchronization During Temporal Dislocation

Bryan-Kinns (Bryan-Kinns and Hamilton 2009; Bryan-Kinns 2004) has identified more meta-information in his studies on mutually engaging music interaction in dislocated setting. He highlights the following properties that are needed to communicate the necessary meta-information for interesting collaboration when both spatial and temporal co-location are not possible.

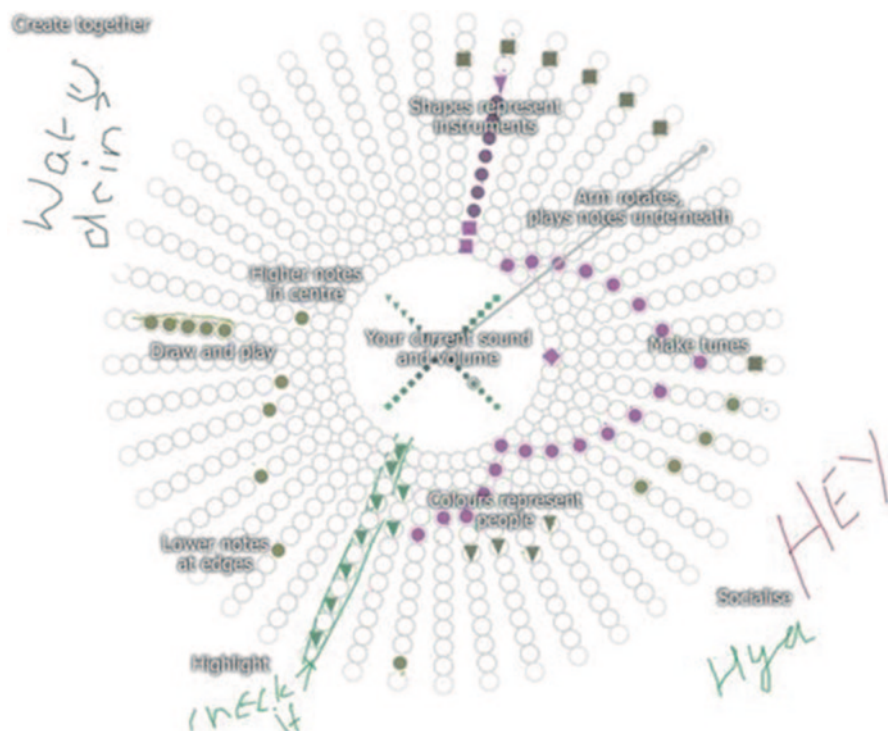
- **The mutual awareness of action**—highlighting new contributions to the joint product and indicating authorship of components
- **Annotation**—being able to communicate in and around a shared product, and being able to refer to parts of the product helps participants engage with each other
- **Shared and consistent representation**—the participants find it easier to understand the state of the joint product and the effects of their contributions if the representations are shared and consistent
- **Mutual modifiability**—editing each other's contributions increases mutual engagement

In the Daisyphone project (Bryan-Kinns 2012) Bryan-Kinns has implemented these properties into the interface, which can be seen in Fig. 4.3. Since the interface is designed with these principles in mind, it allows both temporal and spatial dislocation during a collaboration.

### Coordinating Differentiation

While synchronization is an essential part of joint musical expression, it solely serves the purpose of aligning the expressions of the musicians to each other. In order to make the collaboration dynamic and create new expressions, other tools are used. An established interaction to facilitate this differentiation is the leader/follower dynamic. During the collaboration, a leader will be communicated who can improvise something new into the collaboration. The follower(s) must be able to adapt to the leader's new piece. The leadership role can also be reassigned during the collaboration in order to give multiple collaborators the chance to improvise.

Beaton (Beaton et al. 2010) has established that leadership is communicated over different channels, being the visual, verbal and auditory ones. Reidsma et al. (2008) show that for a virtual music conductor (see Fig. 4.4) to work, it has to take into account what the orchestra is doing (ie. it is a two-way process between leader and follower). Borchers et al. (2004) have shown the same for the inverse case: a virtual



**Fig. 4.3** User-interface of the Daisysophone (Bryan-Kinns 2012) with annotated representation and hand-written textual communication



**Fig. 4.4** Bidirectional interaction with a virtual conductor (Reidsma et al. 2008)

orchestra conducted by a real conductor. This idea is also confirmed by Akkersdijk (2012): the leaders in her experiment used eye contact to monitor whether the follower was actually still following. Akkersdijk also identified several actions that leaders employ to communicate their intentions. Leaders entrain their breathing,

head nods, body movements and arm movements with the rhythm while followers employ mimicry (most in arms, least in head) to show that they are picking up on the proposed rhythm. This is in parallel to the work by Varni et al. (2008) about emotional entrainment to an external rhythm. It could be that while entrainment is used to maintain a rhythm it is also used to change rhythm by the leader. The follower's feedback in the form of mimicry is used by the leader as a measure of how well his improvisation is catching on to his/her partners.

#### 4.4.2 *Performer-Audience Relation*

A different relation is the one between the performer and the audience. While the communication from the musician to the audience is straightforward, being the performance as a whole, the feedback from the audience is a little complicated yet important for the musician to be aware of his/her audience as well. We look at the monologue communication from musician to audience, which we shall call the *Performance* and the dialogue between musician and audience which completes the event, which we will term the *Performance-related interaction*.

##### **Performance**

Besides the created music, several other channels are often employed by the musician in a performance. Munoz (2007) emphasizes the importance of these channels, coining movement as the motor of sound and intention as the impulse of gesture, which leads to the inevitable relation between intentional body movements and musical expression. We found 4 different categories of performance: virtuosity, affective expressions, dance and speech. The Marimba robot (Hoffman and Weinberg 2010) was deliberately designed with only 4 arms to operate a complete Marimba, in contrast with many robotic performance projects that assign an arm to each key on an instrument like the MahaDeviBot (Eigenfeldt and Kapur 2008). This was done to express virtuosity, since the fast movements of a limited number of operating arms to generate the music is perceived as a form of mastery over the instrument. This is further emphasized by Paine (2009) who notes the fast bodily movements as an important factor in the appreciation of a performance by the audience. Musicians employ affective expressions to tell something about and amplify the power of the conveyed emotions in the music. Thompson et al. (2008) show that body movements and facial expressions are often used for this effect. Mobile instruments allow the musician to perform limited dance while playing the instrument. Instruments such as the *Mappe per Affetti Erranti* (Coletta et al. 2008) envision to exploit new interfaces to enable a complete blend of dance and music virtuosity. Verbal expressions are also used to introduce a piece of music or for affective amplification of the music.

## Performance-Related Interaction with the Audience

Tarumi et al. (2012) have informally explored the communication between performer and audience and have found several gestures employed by the audience: hand waving, hand clapping, hand joggling, pushing hands in the air and towel swinging. In addition, the audience also uses words, either short words like “yeah!” and “oh!” or phrases of a song altogether. All these actions share that they are a form of entrainment to the music of the musician, expressing compulsion and engagement. Takahashi et al. (2011) have built hand-clapping robots, placed in front of a streaming performer (performing over an on-line stream) and operated by his remote audience. Visitors of the performance on-line can use buttons on the website to make the robots clap to the musician, enabling hand clapping while being spatially dislocated.

## 4.5 Co-Creative Agents: The Instrument as Agent

Robotic music-making is an emerging field within music technology. We want to look at the collaborative aspects and the interaction with real musicians in the co-creative agent, being an agent with the ability to collaborate with other artificial or human musicians. The preconditions for such an agent are its abilities to perceive and create music and to understand and act in social contexts. We will shortly look at music perception and music creation and agents in a social context (ie. a co-creative agent).

### 4.5.1 *Music Perception*

Before an agent can be part of a musical ensemble, it must, trivially, be able to make music and listen to the music of its partners. Music perception is the topic of perceiving music and has received a significant amount of attention in research, since its applications vary across a wide array of music-related settings. McDermott and Oxenham (2008) have reviewed recent developments in the field. They characterize the main features of music (e.g. pitch, timbre and rhythm) and point out the cultural differences in the way music is perceived. In their study, they elect cognitive and neurosciences as fields where music perception can benefit from. Modelling music perception in a cognitive manner is effective in extracting affective features from music. State-of-the-art examples of systems for the perception of music in a musical agent are Varni’s system for the analysis of interaction in music (Varni et al. 2010), discussed more thoroughly in the next part about social awareness and robotic drummers that are able to collaborate on rhythmic dimensions like the robots built by Weinberg and Driscoll (2006) and Crick et al. (2006).

Hoffmann and Weinberg's *Shimon* (2010) reduces the problem of harmony perception to simple pitch detection. Instead of working with a note representation of what it hears, it just listens to the pitch of the music and moves its robotic arms to the approximate location of corresponding pitch locations on the Marimba. The anticipatory system in the Marimba furthermore builds patterns of how its partner progresses through harmony to anticipate where the arms should be, making it possible to play along without delay.

### 4.5.2 Music Creation

Meta-creation involves using tools and techniques from artificial intelligence, artificial life, and machine learning, themselves often inspired by cognitive and life sciences, to create music. Musical meta-creation suggests exciting new opportunities to enter creative music making: discovery and exploration of novel musical styles and content, collaboration between human performers and creative software “partners”, and design of systems in gaming and entertainment that dynamically generate or modify music. We have identified three general approaches in musical agents for meta-creation, being the model-driven, the data-driven and the cognitive approaches. We also look at a specific application of meta-creation for improvisation: continuation.

The **model-driven** approach models music in terms of aesthetic and organized rhythm and harmony. This model of music is then used by the agent to produce music that fit the model. Eigenfeldt et al. have developed Harmonic Progression (Eigenfeldt 2009). Another example is the Kinetic engine, a real-time generative system (Eigenfeldt 2010) which has been used in different contexts, amongst which improvising ensembles (e.g. MahaDeviBot (Eigenfeldt and Kapur 2008)).

Another approach to music creation is the **data-driven** approach. This strategy generates music from a pool of existing compositions. The idea is implemented in MusicDb (Maxwell and Eigenfeldt 2008), which consists of a music information database and an accompanying query system, which both form a streamlined algorithmic composition engine.

The **cognitive learning** method models a music creating agent to imitate human-like musical creativity. It is concerned with the learning process: what features of music do we store and how do we store them in our minds? The hierarchical memory theory (Lerdahl and Jackendoff 1996) defines a hierarchical storage, while the Long and Short Term Memory approach (Eck and Schmidhuber 2002) emphasizes the importance of temporal features in this hierarchy. Both theories are applied in the Hierarchical Sequential Memory for Music (Maxwell et al. 2009).

The Sony Continuator (Pachet 2002) is an example of **continuation**, a special form of meta-creation in which a machine takes up the music after the human player stops, while maintaining the style initiated by the human performer. By employing Monte Carlo Markov chains (Gamerman and Lopes 2006), the Continuator learns

the music style of a performer and then continues in the same style, providing a partner during improvisation.

### 4.5.3 *Social Awareness and Behavior*

A part of co-creation that is especially interesting in the light of mediated interaction is social behavior and awareness in co-creative agents. A co-creative agent must be able to understand and act out the multi-modal communication described in Sect. 4. We discuss a few systems on social awareness and then proceed to highlight a few projects on social behavior in collaborative settings. We will also shortly visit choreography as a special category of social behavior.

#### **Social Awareness**

For an autonomous musical agent to function in a collaborative setting, it needs to be aware of that setting. A system for the analysis of social interaction in music has been developed by Varni et al. (2010). The key factors of their system are driving interpersonal synchronization of participants in a music ensemble, the identification of roles (e.g., leadership, hierarchy), the general principles of how individuals influence each other and the factors that drive group cohesion and the sense of shared meaning. This system shows promises for the development of social awareness in collaborative agents for musical expression.

#### **Social Behavior**

Social behavior is the acting out of the social identity of the agent in a collaboration. The agent must be able to act accordingly to social rules. Communication between musicians was studied in Sect. 4. This information is useful to implement social agents.

Shimon (Hoffman and Weinberg 2010) was already provided as an example of a social musical collaborator. It contains several interaction modules to make music in a social way. Its *call-and-response* system responds to a musical phrase with a chord sequence. Its *opportunistic overlay improvisation* tries to anticipate upcoming chords by its partners and play notes within that chord. The *rhythmic phrase-matching improvisation* is a decaying-history probability distribution for rhythm, classifying which rhythm to use to stay synchronized with its partner.

If we split up roles in *following* and *leading*, knowledge from the related field of conductor simulation can be used as a starting point. Borchers et al. (2004) have developed the Personal Orchestra, a conducting environment in which the user can conduct a virtual orchestra. The conducting behavior of the user is translated into musical parameters (e.g. tempo) and the virtual orchestra adjusts how they play

a (predefined) music piece according to those parameters. Combining the extraction of parameters with Eigenfeldt & Pasquier's Harmonic Progression (discussed above), we can let a model-driven creative agent be influenced by the user directly. The MIROR project (MIROR 2012) with its focus on reflexive music generation can also serve well in such a situation. Their installation is able to react upon the user's input without interrupting the music.

The counterpart of a following role is the leading role. Beaton et al., in their study that compares experienced musicians with inexperienced people (Beaton et al. 2010) in the context of human-computer musical collaboration, report that inexperienced musicians perceived a leading agent as helpful to guide them through the process of making music. This contrasts with their findings on experienced musicians in the same setting: they tend to prefer a following musician. This means that when the intended user category of an autonomous music creating agent is the non-musician, a leading role is favorable. To implement leading properties, again we can look at research on conducting systems. In this case, we look at a virtual conductor that leads a human orchestra made by Reidsma et al. (Raffa 2011). It is pointed out that to successfully lead in a musical setting, one needs to communicate intentions and work well together with the human musicians. This is why special attention is paid to the subtle dynamic of leading and following when fulfilling the role of a leader.

In Shimon, an artificial head is used to communicate leadership: when the robot looks at the musician, it wants him to improvise (and will follow the musician). If it looks at its Marimba, it will improvise itself (lead).

## Choreography

Choreography is especially interesting when the robotic musician is performing for an audience. As we have shown in Sect. 4, the expression of virtuosity, affect and even dance through the non-verbal channel are important for an interesting performance. It can also serve as a way of communication, using realistic movements when playing an instrument to give other musicians a way to anticipate the robot's musical phrases. Liveness has also been noted in several studies as an important factor for a performing robot, both for an audience and for its partners (Hoffman and Breazeal 2008). A choreographic system is needed for the robot to be interesting in a performance and pleasurable as a music partner.

## 4.6 Discussion

In this survey, we have established a broad view of the state-of-the-art in instruments for music expression. Instrument design principles have been categorized in the tool, media and agents paradigms as well as into different user categories, of which the main distinction is between musicians and non-musicians.



When viewing instruments as tool, the instrument is said to be a thing that the human user can use to interact with the physical world. It is important to consider the user groups that the instrument is being designed for. The interface can be tangible, which is a preferred interface for musicians as tangibility allows natural feedback to occur in the form of haptics. Tangibility is preceded by the touch-interface and followed up by haptics in terms of interaction sophistication. Other interfaces can be gesture-based, imagery-based or iconic and touch-oriented. The type of interface can have a strong influence on the control mechanism. Whereas tangible interfaces provide a robust control mechanism that enables direct control, it is very hard to have direct control with gesture-based controls, although there have been successful attempts and studies are ongoing on the topic of gesture sonification. Not only the interface but the exterior design is also an important determinant for the reception of an instrument. The use of computer interfaces can suffer the lack of allure that is present in performances on traditional instruments. When designing for a specific type of user, the exterior design should also be adapted towards the expectation of that user group.

Instruments can be designed as media, which means that they transcend the concept of a tool. Media enable the channeling of communication between different humans as part of their design and thus enable interesting new forms of musical collaboration. We distinguished between spatial and temporal co-location in a collaboration. Usually, musicians are both spatially and temporally co-located when performing music with tools. With media, we are able to remove one or both forms of co-location to conceive new forms of musical collaboration. When designing a spatially de-located instrument, it is important to make sure that the communication that is present otherwise is substituted in some manner. A definer of success is to provide a means for the user to distinguish between the tunes made by the different collaboration partners. Definers of success in temporally de-located collaborations are the mutual awareness of action, annotation capabilities, shared and consistent representations and mutual modifiability. The possibilities of musical instruments as media are huge, promising flexibility, new forms of social interaction and user-generated artistic content in public areas.

When designing agents for musical expression, a main issue is the design of computational creativity. There are three main categories of computational musical creativity. The first is the model-driven approach, in which the designer can model the generation of aesthetically pleasing music and feed that model to the agent. Main studies have been concerned with the modelling of rhythm and harmonics. Secondly, we have the data-driven approach, which is the art of building an agent that is able to combine different existing musical pieces in seamless transitions in an interesting way. The last approach is cognitive learning, in which we try to model an agent's musical creativity learning process as humans do so. Key concepts are long and short term memory, hierarchical memory and neural network methods.

We are not only interested in robotic music authorship, which is the area of computational creativity, but also in the possibilities of a virtual musician augmenting a human in his musical expressions. This is when an agent can serve as a smart instrument. Next to computational creativity, one of the things we also have to account

for are social awareness, which primarily consists of role recognition and role execution within a musical ensemble. There are promising systems for the analysis of social context within musical ensembles. Theory on the execution of musical roles can be lent from the domain of musical conductor research. Studies have shown that non-experienced users favor different social behavior from the agent than experienced musicians. The last important piece to enable autonomous agents as part of an instrument is music perception. The agent must have the ability to experience the music of the human musician and for this, we need to implement methods of music perception. It has been suggested that the fields of cognitive and neurosciences can teach us about the cognitive methods of music perception.

In this chapter, we discussed various aspects of coordination in technology enhanced musical expression. The next step would be to better understand the interactive element of micro-coordination in collaborative musical expression, to ultimately build prototypes that can either substitute traditional communication channels or augment the interactive landscape in micro-coordination. Such research would target questions such as: What actions exist in the interactive landscape in collaborative musical expression (continuing on findings presented in this survey)? How can technology be used without disrupting this landscape? How can we use technology to enrich these interactions? We are looking at a field that, although already well-explored by many researchers, is still full of exciting possibilities for radically new interactive experiences.

## References

- Akkersdijk S (2012) Synchronized clapping: a two-way synchronized process. *Capita Selecta paper*
- Barbosa Á (2008) Ten-hand piano: a networked music installation. In: *Proceedings of the 2008 Conference on New Interfaces for Musical Expression, NIME '08, Genova, Italy*, pp 9–12
- Barbosa Á, Kaltenbrunner M (2002) Public sound objects: a shared musical space on the web. In: *Proceedings of the First International Symposium on Cyber Worlds, CW '02, Washington, DC, USA*, pp 9–11
- Beaton B, Harrison S, Tatar D (2010) Digital drumming: a study of co-located, highly coordinated, dyadic collaboration. In: *Proceedings of the 28th international conference on Human factors in computing systems, CHI '10, New York, USA*, pp 1417–1426
- Beilharz K (2011) Tele-touch embodied controllers: posthuman gestural interaction in music performance. *Social Semiotics*, Vol 21, issue 4. (Published Online), pp 547–568
- Benthien C (2002) *Skin: on the cultural border between self and the world*. Columbia University Press. New York, USA
- Bianchi-Berthouze N (2013) Understanding the role of body movement in player engagement. *Hum Comput Int* 28(1):40–75. (Published Online)
- Borchers J, Lee E, Samminger W, Muhlhauser M (2004) Personal orchestra: a real-time audio/video system for interactive conducting. *ACM Multimedia Systems Journal Special Issue on Multimedia Software Engineering*, vol 9, issue 5, San Jose, USA, pp 458–465
- Bryan-Kinns N (2004) Daisyphone: The design and impact of a novel environment for remote group music improvisation. In *Proceedings of the 5th Conference on Designing Interactive Systems: processes, practices, methods, and techniques, DIS 2004*, pp 135–144

- Bryan-Kinns N (2012) Mutual engagement in social music making. In: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, LNICST 78, pp 260–266. (Published Online)
- Bryan-Kinns N, Hamilton F (2009) Identifying mutual engagement. *Behav Inform Tech* 31(2):101–125
- Chew YCD, Caspary E (2011) MusEEGk: a brain computer musical interface. In: CHI '11 extended abstracts on human factors in computing systems. New York, NY, USA: ACM, pp 1417–1422. (A BCI example is MusEEGk)
- Choi I, Bargar R (2011) A playable evolutionary interface for performance and social engagement. *INTETAIN*, vol 78. Genova, Italy, pp 170–182
- Coletta P, Mazzarino B, Camurri A, Canepa C, Volpe G (2008) Mappe per Affetti Erranti: a multimodal system for social active listening and expressive performance. In: Proceedings of the 2008 Conference on New Interfaces for Musical Expression, NIME '08, Genova, Italy, pp 134–139
- Crick C, Munz M, Scassellati B (2006) Synchronization in social tasks: Robotic drumming. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication. ROMAN, Vol 15, Hatfield, UK, pp 97–102
- De Boeck C (2011) Staalhemel: responsive environment for brainwaves. Available: <http://www.staalhemel.com/>
- Eck D, Schmidhuber J (2002) Finding Temporal Structure in Music: Blues Improvisation with LSTM Recurrent Networks. In: Neural Networks for Signal Processing XII, vol 12. Martigny, Valais, Switzerland, pp 747–756
- Eigenfeldt A (2009) A realtime generative music system using autonomous melody, harmony, and rhythm agents. In: XIII Internationale Conference on Generative Arts, Milan, Italy
- Eigenfeldt A (2010) Realtime generation of harmonic progressions using controlled markov selection. In: Proceedings of the First International Conference on Computational Creativity, ICC 2010. Lisbon, Portugal
- Eigenfeldt A, Kapur A (2008) An agent-based system for robotic musical performance. In: Proceedings of the 2008 Conference on New Interfaces for Musical Expression, NIME '08, Genova, Italy, pp 144–149
- Fikkert FW, Hakvoort MC, van der Vet PE, Nijholt A (2009) Feel sound: interactive acoustic music making. In: Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACE2009, New York, pp 449–449
- Gamerman D, Lopes HF (2006) Markov chain Monte Carlo: stochastic simulation for Bayesian inference, vol 68. Chapman & Hall/CRC
- Georgeff M, Pell B, Pollack M, Tambe M, Wooldridge M (1998) The belief-desire-intention model of agency. *Intelligent Agents V: Agents Theories, Architectures, and Languages*. Paris, pp 1–10
- Goina M, Polotti P (2008) Elementary gestalts for gesture sonification. In: Proceedings of the 2008 Conference on New Interfaces for Musical Expression, NIME '08, Genova, Italy, pp 150–153
- Harmonix Music Systems (2005) Guitar hero video game series. DVD by RedOctane
- Hoffman G, Breazeal C (2006) Robotic partners' bodies and minds: an embodied approach to fluid human-robot collaboration. *Cognitive Robotics*
- Hoffman G, Breazeal C (2008) Anticipatory perceptual simulation for human-robot joint practice: theory and application study. In: Proceedings of the 23rd national conference on Artificial intelligence, Chicago, IL, USA, pp 1357–1362
- Hoffman G, Weinberg G (2010) Shimon: an interactive improvisational robotic marimba player. In: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems. ACM, pp 3097–3102
- Hoffman G, Kubat R, Breazeal C (2008) A hybrid control system for puppeteering a live robotic stage actor. In: Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on. IEEE, pp 354–359
- Ihde D (1986) *Experimental phenomenology: an introduction*. SUNY Press
- Jorda S, Alonso M (2006) Mary had a little scoreTable\* or the reacTable\* goes melodic. In: Proceedings of the 2006 conference on New interfaces for musical expression. IRCAM Centre Pompidou, Paris, France, pp 208–211

- Jorda S, Geiger G, Alonso M, Kaltenbrunne M (2007) The reactable exploring the synergy between live music performance and tabletop tangible interfaces. In: the Proceedings of the 1st Conference on Tangible and Embedded Interaction, New York, NY, USA
- Kerckhove D (1993) Touch versus vision: Ästhetik neuer technologien. Die Aktualität des Ästhetischen, Germany, pp 137–168
- Le Groux S, Manzolli J, Verschure P (2010) Disembodied and collaborative musical interaction in the multimodal brain orchestra. In: Proceedings of the International Conference on New Interfaces for Musical Expression, pp 309–314
- Lee JS, Tatar D, Harrison S (2012) Micro-coordination: because we did not already learn everything we need to know about working with others in kindergarten. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, CSCW '12, New York, NY, USA, pp 1135–1144
- Lerdahl F, Jackendoff R (1996) A generative theory of tonal music. The MIT Press, Cambridge
- Leslie G, Mullen T (2011) MoodMixer: EEG-based collaborative sonification. Proceedings of the International Conference on New Interfaces for Musical Expression, 30 May–1 June 2011. Oslo, Norway, pp 296–299.
- Levin G (2006) The table is the score: an augmented-reality interface for real-time, tangible, spectrographic performance. In: Proceedings of the International Computer Music Conference (ICMC' 06), New Orleans, USA
- Maxwell J, Pasquier P, Eigenfeldt A (2009) Hierarchical sequential memory for music: a cognitive model. In: the Proceedings of the 10th International Society for Music Information Retrieval Conference. ISMIR 2009, Kobe International Conference Center, Kobe, pp 429–434
- Maxwell JB, Eigenfeldt A (2008) The MusicDB: a music database query system for recombination-based composition in Max/MSP. In: the Proceedings of the International Computer Music Conference, ICMC2008. (Published Online)
- McDermot J, Oxenham AJ (2008) Music perception, pitch and the auditory system. In: the Proceedings of the Current Opinion in Neurobiology 18(4):452–463 (geen locatie!)
- Miller K (2009) schizophonic performance: guitar hero, rock band, and virtual virtuosity. *J Soc Am Music* 3(4):395–429
- Miranda E (2006) Brain-computer music interface for composition and performance. *Int J Disabil Hum Develop* 5(2):119–126
- MIRROR website (2012) Musical interaction relying on reflection. <http://www.mirrorproject.eu>
- Munoz EE (2007 Nov) When gesture sounds: bodily significance in musical performance. In Conference proceedings from the International Symposium on Performance Science, ISPS2007. Porto, Portugal, pp 55–60
- Nijholt A, Tan DS, Allison BZ, Millán José del R, Graimann B (2008) Brain-computer interfaces for HCI and games. *CHI Extended Abstracts* 2008:3925–3928
- Pachet F (2002 Sep) The continuator: musical interaction with style. In: Proceedings of the International Computer Music Conference. ICMA2002. Gothenburg, Sweden, pp 211–218
- Paine G (2009 Aug) Towards unified design guidelines for new interfaces for musical expression. *Organ Sound* 14(2):142–155. (Published Online)
- Patten J, Recht B, Ishii H (2002) Audiopad: a tag-based interface for musical performance. In: Proceedings of the 2002 conference on New interfaces for musical expression. National University of Singapore, Singapore, pp 1–6
- Poupyrev I, Berry R, Kurumisawa J, Nakao K, Billinghamurst M, Airola C, Kato H, Yonezawa T, Baldwin L (2000) Augmented groove: collaborative jamming in augmented reality. In: ACM SIGGRAPH2000 Conference Abstracts and Applications, New Orleans, USA, p 77
- Raffa R (2011) Rhythmsynthesis: visual music instrument. In: Proceedings of the 8th ACM conference on Creativity and cognition. ACM, pp 415–416
- Reidsma D, Nijholt A, Bos P (2008 Dec) Temporal interaction between an artificial orchestra conductor and human musicians. *Computers in Entertainment* 6(4):1–22
- Rosenbaum E (2008) Melodymorph: a reconfigurable musical instrument. In: Proceedings of the 8th International Conference on New Interfaces for Musical Expression. NIME2008. Genova, Italy, pp 445–447

- Sänger J, Müller V, Lindenberger U (2012) Intra- and interbrain synchronization and network properties when playing guitar in duets. *Front Hum Neurosci.* 2012; 6: 312. Published online 2012 November 29. Prepublished online 2012 July 6. doi: 10.3389/fnhum.2012.00312
- Takahashi M, Kumon Y, Takeda S, Inami M (2011) Remote hand clapping transmission using hand clapping machines on live video streaming. In: *Proceedings of Entertainment Computing*
- Tarumi H, Akazawa K, Ono M, Kagawa E, Hayashi T, Yaegashi R (2012) Awareness support for remote music performance. *Advances in Computer Entertainment, ACE2012, Nepal*, pp 573–576
- Tatar D (2012) Human-computer interaction seminar. University Lecture, Stanford
- Teitelbaum R (1976) *In tune: some early experiments in biofeedback music (1966–1974)*. Vancouver, BC, Canada: Aesthetic Research Center of Canada Publications
- Thompson WF, Russo FA, Quinto L (2008) Audio-visual integration of emotional cues in song. *Cogn Emot* 22(8):1457–1470, Published Online
- Varni G, Camurri A, Coletta P, Volpe G (2008) Emotional entrainment in music performance. In: *Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition. FG2008*. Amsterdam, The Netherlands, pp 1–5
- Varni G, Volpe G, Camurri A (2010) A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. *IEEE Transactions of Multimedia*, vol 12, pp 576–590. (Published Online)
- Verbeek P-P (2005) *What things do: philosophical reflections on technology, agency and design*. Science and Engineering Ethics, Pennsylvania State University Press
- Verplank Bill (2011) Keynote on motors and music. Presented on Chi Spark 2011 as the second keynote
- Verplank Bill, Gurevich M, Mathews M (2002) The plank: designing a simple haptic controller. In: *Proceedings of the 2002 conference on New interfaces for musical expression. NIME2002*. Singapore, pp 1–4
- Ward KPN, O'Modhrain S (2008) A study of two thereminists: towards movement informed instrument design. In: *Proceedings of the 2008 Conference on New Interfaces for Musical Expression. NIME '08*. Genova, Italy, pp 117–121
- Weinberg G, Driscoll S (2006) Robot-human interaction with an anthropomorphic percussionist. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems. CHI2006*. New York, USA, pp 1229–1232