An Approach to Collaborative Music Composition

Niklas Klügel, Marc René Frieß, Georg Groh
Technische Universität München
Boltzmannstraße 3
85748 Garching, Germany
{kluegel|friess|grohg}@in.tum.de

Florian Echtler
Munich University of Applied Sciences
Lothstraße 64
80335 München, Germany
florian.echtler@hm.edu

ABSTRACT

This paper provides a discussion of how the electronic, solely IT based composition and performance of electronic music can be supported in realtime with a collaborative application on a tabletop interface, mediating between single-user style music composition tools and co-located collaborative music improvisation. After having elaborated on the theoretical backgrounds of prerequisites of co-located collaborative tabletop applications as well as the common paradigms in music composition/notation, we will review related work on novel IT approaches to music composition and improvisation. Subsequently, we will present our prototypical implementation and the results.

Keywords

Tabletop Interface, Collaborative Music Composition, Creativity Support

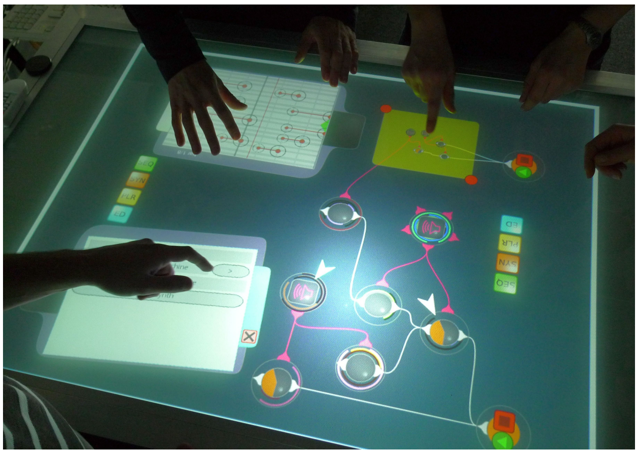

Figure 1: The Application in Use

1. INTRODUCTION

The adoption of electronic devices into music composition, performance and perception since the middle of the last century culminated in electronic music being part of today's popular culture. At the same time, new forms of music composition were introduced to a larger group of people, thereby gradually popularizing IT supported music composition. With this paper we want demonstrate how music composition can be transformed to become a more situative and collaborative process. We emphasize the discursive exploration of expression and exchange of musical ideas, thereby focusing on experts in the field of electronic music composition (cp. [2]). In this context it is worthwhile stating that we restrict ourselves to "electronic composition of electronic music", where "electronic composition" means that musical elements are solely manipulated with computers and "electronic music" that all sounds are synthesized by a computer as well.

Unfortunately, most support tools for composing electronic music collaboratively focus on distributed and asynchronous collaboration. The attention on distribution and single-user user-interfaces poses a problem in view of designing truly collaborative applications as "*the user's concentration has to shift away from the group and towards the computer*" [9]. This contrasts to the importance of the social context and social interactions for creative tasks [4]. Face-to-face interaction in co-located settings has advantages over electronically mediated interaction such as the visibility of action and the ability to use verbal and non-verbal communication. As novel means of IT-support for performing music face-to-face, tabletop based applications such as the reacTable [13] are promising, as they particularly support this kind of social context. While in the single user case of creatively manipulating music on a stand-alone PC, every action can be arbitrarily stored, undone, resumed and reflected on, the improvisational social case lacks these possibilities for obvious reasons. Considering these aspects, the research question to be answered within this paper is:

> *How can the advantages of co-located collaborative musical improvisation and single-user oriented applications for systematic and iterative music composition be combined in tabletop-based system that synchronously, collaboratively and in a co-located way allows for the composition and performance of music?*

In order to devise an application suited to answer this question, at first requirements for applications to support (group) collaboration on such devices in general have to be discussed. Afterwards we will discuss some paradigms of music composition in general and will review related work in regard to distributed as well as co-located IT systems for music composition and performance. Finally the application will be described in detail.

2. GROUP COLLABORATION

Multi-touch tabletop devices feature several characteristics that support collaboration as they feature intuitive methods of shared and parallel input. Gestural (and tangible) interfaces exploit the human kinesthetic capabilities gained from life-long interaction with the physical environment [10]. This results in both a rich vocabulary for Human Computer Interaction and the reduction of cognitive overhead for tactile interaction [5, 11]. These fundamental properties for tabletop devices overlap in their consequence with properties desired and essential to collaboration: Barriers in the computer mediated interaction between collaborators are alleviated and the immediate joint use of interaction is fostered. While these characteristics provide a basis for enabling situated collaboration, a practical realization is bound to the domain of application and design principles solving high-level issues. In this regard, the following questions arise: How to support the complex dynamics of collaboration? Which constructs in the application support a high degree of synchronicity in the collaboration?

While it is all but impossible to grasp collaboration holistically, there are several aspects in collaboration that have been identified to play a key-role for realizing applications. In the field of Computer Supported Collaborative Work (CSCW) those are: **group awareness**, **group articulation** and **tailorability** [3, 6]. **Group awareness** describes the ability of individuals in the group to share and gain mutual knowledge about the activities performed. Hence, a good sense of group awareness helps collaborators to coordinate activities and to access shared resources by simplifying communication. **Group articulation** refers to the ability to partition the group task into units that can be combined, substituted and reintegrated in the group task, while **tailorability** proclaims the ability for individual adoption of technology (in unintended ways).

Concerning tabletop interfaces, various forms of **coupling** can be identified [18]. Coupling describes the different styles of the mediated collaboration around the tabletop device and as such it results in a steady flux of the group configuration. Supporting such group transitions is the key to permitting the dynamics of collaboration to happen, since they reflect a natural style of communication and interaction between collaborators. The spatial usage of the tabletop environment for personal and group tasks (**territoriality**) is also relevant. It helps to coordinate actions with artifacts or with collaborators on the tabletop [17] akin to the human division of space in the physical reality. In general, the parallel interaction with the application are very important. On one hand, they foster group awareness and group articulation since centralized controls may disrupt the workflow [14, 10]. On the other hand they make contributions for all group members more democratic because of the distributed right to control artifacts. In this way, individual users are not prioritized and thresholds for collaborating are lowered [7]. These aspects will serve as the basic requirements for our application design. How we explicitly address these issues like territoriality and coupling will be discussed in section 5. Before, the application field of music composition as well as related work needs to be regarded.

3. COMMON PARADIGMS IN MUSIC COMPOSITION & NOTATION

It is important to briefly discuss some basic concepts of music composition, notation and the relevant prerequisites for collaboration in the domain of this application to further frame the requirements for the design and its musical applicability.

A key aspect for composition is to establish a notational system to express the comprising musical events. Since the sonification of the composition is performed by the computer, its interpretation and execution by a performer is omitted. Thus, the control over the perceptual dimensions such as timbre is part of the notation and thereby the composition. Therefore, for electronic music the notation serves as a blueprint for the stream of sound - tailored to the specific requirements of musical expression for a particular piece. But especially in the context of collaborative use of the notation, the problem is to create a musically expressive set of notational symbols whose meaning and effect is plain to all participants. Furthermore the notation must support collaboration in a concurrent and simultaneous way.

Traditional music admits various organizing principles, both hierarchical and non-hierarchical, making the communication of musical ideas a knowledge representation problem. In a collaborative setting, it is essential for musical expression to support such organizing principles and to make the compositional structure modifiable and, for the dialogue, holistically comprehensible. But this is conceptually difficult to realize for a shared tabletop interface. It is caused by the representation (see section 5.3 - User Interface) and the rigid structure of a (traditional) score for the collaborative context. In many cases, the development of musical ideas into the resulting final composition is inherently bound to properties that have been chosen at the beginning. Insofar the process of composing itself is constrained by assumptions taken gradually but on a global level (e.g. the instrumentation of a piece, tempo changes) while the locality, the independence of (sonic and temporal) characteristics for outlining a musical idea is not supported. An application for collaborative music composition therefore has to allow the creation of meta-/intermediate arrangements of a composition - a conceptual unification of creating arbitrary musical sketches (separate arrangements) that can coexist and be gradually shaped into a final composition/arrangement. In view of enriching the overall workflow, solving this problem of conceptional unification is the key point in a collaborative environment. In spite of the high degree of complexity inherent in the process of composing, several application supporting electronic performance as well as composition of music have already emerged. Some tabletop applications will be presented next.

4. RELATED WORK

Tabletop interfaces reflect one end of the spectrum of synchronicity for supported group interaction. Interaction is carried out physically and locally using a single shared workspace. In this regard, many tabletop music applications share goals with ours. Their field of application widely ranges from multi-user instruments (e.g. reactTable [13]) to probabilistic composing environments (e.g. Xenakis [1]). Some of them use physical objects (**tangibles**) that represent artifacts or functionality of the application. As they act as a bridge to the digital representation of an object they make use of our kinesthetic and spatial intelligence for interacting with the application. Thus, tangibles are by concept excelling implementations of the direct manipulation metaphor. Group awareness is assisted as all information that is conceived for interaction with the application is physically apparent, theoretically with all hidden states removed.

ReactTable [13] is a prominent example of such applications. It is a multi-user instrument portraying a modular synthesizer where the signal processing can be altered by gestural and physical interaction using tangibles. The signal flow is represented by the topological relationship between artifacts. As an extension to the original reactTable, scoreTable [12] allows the composition of looped phrases using tangibles while following a spatial approach to composition: the positions where objects are placed determine their values and their functionalities.

Conceptually, tangibles pose problems that are disparate with our requirements. In essence, the hard object relationship (virtual to physical) defeats arbitrary application of tangibles. First, multi-modality is difficult to express with physical objects. In a broader sense, it is challenging to superimpose concepts like context sensitivity on physical objects for an application. But they are essential if one wants to portray and rely on multiple object relationships. Second, it is difficult to represent hierarchical structures with physical objects in a one-to-one mapping. Furthermore, the complexity of a composition would be limited to the number and spatial properties of physical objects. Third, and most importantly, the common use of spatial relationships between tangibles to control the application is problematic for the group articulation as the group task cannot be arbitrarily partitioned into units and shared.

5. THE APPLICATION

We further frame the requirements for our application in regard to the user interface and the underlying data model:

5.1 Requirements

On the technical side, to support the dynamics of the collaborative interaction, the data model and operations have to fulfil these properties in terms of modifiability:

- Concurrency
- Radical, non-linear changes
- Real-time capabilities

For these properties to be perceived and utilized by users, the user interface has to commit to:

- Provide instantaneous feedback on the changes
- Provide consistent feedback on the state of the application / compositional structure and parameters of the synthesis
- Apply the aforementioned design paradigms and rough guidelines for supporting the group collaboration

To support the various aspects of Coupling and Territoriality the operational controls have to be independent of positioning. Furthermore, they should provide concurrent use without evoking ambiguous states and help reducing mode errors [16]. The latter is important, since the less the user's attention has to be drawn to controlling the application appropriately, the more the group communication and awareness is fostered.
We decided to split the user interface into two levels: the first level is for creating, modifying and controlling the compositional structure and the second one is for changing the properties of the elements in it.

5.2 Structures for Composition & Synthesis

The first level represents the compositional structure as a directed acyclic graph. It can be modified in real-time by the user interface. The graph is composed of two forests for the respective functional domains: the temporal order of musical events (sequence graph) and the audio synthesis (synthesis graph). The nodes in the first forest are sequences (cp. figure 2, 4a) which are usually short musical phrases, with the exception of the root nodes. Sequences can contain either arbitrary or chromatic control data (e.g. notes) for the synthesis. Here, the edges denote the succession of sequences; this constitutes the temporal order of an (meta-) arrangement. Each such arrangement is therefore represented as a single tree. The second forest is used to map the musical events of sequences to parameters of synthesizers. Therefore nodes in this representation are either sequences or synthesizers (cp. figure 2, 4b). Edges describe the flow of control information such as a sequence controlling pitch or frequency of a filter of a synthesizer. Parallel edges are allowed if a sequence is to control more than a single parameter of a synthesizer. The two forests are merged visually (sequence nodes exist in both forests so only one is visualized), as a result, the linear notion and depiction of instrument staves and a single timeline is circumvented. On the whole, this concept can be seen as a subset of a data-flow language that allows for graphical patching such as Pure-Data [15], although the application follows a different architecture internally.

Sequencing & Control of playback: A sequence is the atomic entity keeping track of musical events. It allows to insert and remove musical events and maintains their local temporal order. That is, all information regarding the timing of musical events is treated as relative to the beginning of the sequence (start and length can be changed). The type of succession expressed by linking nodes with edges can form either a sequential chain or branches of parallel sequences - this structure corresponds to a tree. The construction of parallel sequence chains equals several instrument staves with phrases - similar to traditional notation. The root node of a tree is a special object to control the playback of a sequence tree independently from others (cp. figure 2, 4c). It not only controls the immediate start and stop of the playback but also enables its synchronization to the global tempo and the looped playback.

5.3 User Interface

The visualizations of sequence and synthesizer nodes have additional graphical indicators : for sequences the current local playback position and for synthesizer the current signal volume generated. Naturally, edges are indicated by a line segment. They can be created by performing a dragging motion from one node to another one. The shared user interface provides areas on the borders, each for the specific type of nodes. Dragging from one of them to the desired position creates motion from one of them (cp. figure 2, 3) to the desired position. With the facility for users to arrange the visualization of the graphs freely and the techniques described thus far, the following conceptual gains for collaboration over traditional DAWs have been established:

- Building blocks of the composition can be freely moved around, grouped and interacted with regardless of the virtual position of artifacts
- Multiple arrangements/sketches of musical ideas are supported on the same interface. In fact there is no notion of a primary arrangement.
- Arrangements can be freely combined and rearranged.
- Changes on the compositional structure are immediately reflected visually and sonically

Manipulating Properties: The second level for interaction is used to change properties of nodes such as the type of the synthesizer or the musical events in a sequence. Chromatic musical events are manipulated using the piano-roll metaphor (common in modern DAWs). Setting properties of nodes in the graph structure is realized with the help of two elements in the user interface: the first one is for selecting the node (cp. figure 2, 1) whose properties are to be changed and the second one for visualizing and manipulating them (cp. figure 2, 2). The EditorView is similar to a window in the WIMP paradigm which floats above the compositional structure, but it can be rotated and translated arbitrarily to match the user's orientation. It also provides facilities to pan and zoom the visualization and to close the view. The Selector can be dragged onto a node in the graph structure. Synchronous to this, the contents in the EditorView are updated showing an interface to manipulate the selected node's properties. An arbitrary number of these Selector/EditorView pairs can be instantiated to edit the properties of multiple graph nodes concurrently. An integral feature is that the content in an EditorView again can be shared in a synchronous way, making the properties of a single item accessible for editing from several EditorViews. This allows users to concurrently interact and gain mutual insight in their doing, thus fostering group awareness.

Gestural Control: Most multi-touch capable tabletop surfaces do not associate touch points to a to a specific user. Here, collaborative gestural input can be prone to create ambiguous states in conjunction with multi-touch gestures. As a result, it was decided to utilize primarily single-touch gestures for interaction with objects that are intended to be shared. Multi-touch gestures were only applied for controlling the EditorView (to change the view and orientation). The limited amount of gestures is sufficient for the provided functionality while helping habituation (reducing cognitive overhead / fostering group awareness).

6. CONCLUSION

The control over the application is independent of the orientation or position of artifacts in the user interface. Ergo, the flexible and dynamic coupling of users while performing the group task is supported. This is achieved by avoiding spatial relations to express temporal and tonal dependencies of the composition with help of the graph visualization and is further extended with spatially independent and context-sensitive editors. Since all aspects about the modification and, not to be disregarded, the playback of the shared compositional structures can be controlled in real-time, the

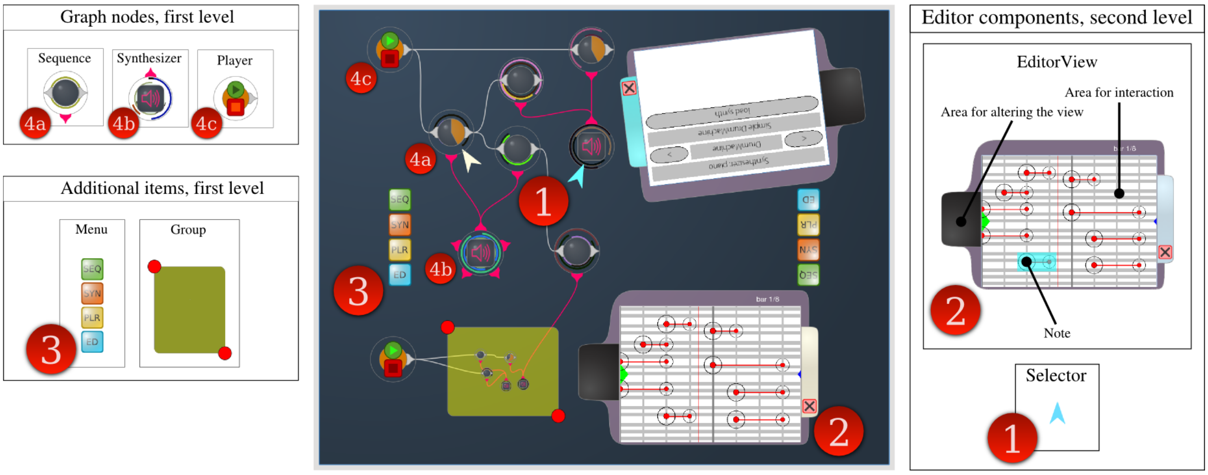

Figure 2: Overview of the Application's User Interface and Control Objects

discourse of musical ideas can be shifted from a verbally dominant one to a hands-on approach. Building blocks of compositional structures can be combined with instant sonic feedback. This simplifies the creative exploratory discourse while it blurs the line between composing collaboratively and improvising collectively:

- Users are encouraged to sketch and shape their musical ideas in real-time, synchronized with those that have been created or are in process of creation by others.
- The playback of disjoint arrangements can be triggered ad-lib by the users to fabricate new and derivative arrangements promptly.

In this paper an analysis of interface principles for collaborative co-located applications has been conducted. It has been shown that the dynamics and complexity of collaborative work imply numerous constraints on the design of our application. After the discussion of related work, the accompanied findings have been used as guidelines for the overall design of the prototype. In this way the implementation is rooted in user and collaboration centric design as opposed to building an interface on top of a preexisting application or framework to convey the inherent functionality. A demonstration video of the application can be seen in [8].

7. REFERENCES

[1] M. Bischof, B. Conradi, P. Lachenmaier, K. Linde, M. Meier, P. Pötzl, and E. André. Xenakis: combining tangible interaction with probability-based musical composition. In *TEI '08: Proc. of the 2nd Int. Conf. on Tangible and embedded interaction*, pages 121–124, New York, NY, USA, 2008. ACM.

[2] T. Blaine and S. S. Fels. Contexts of collaborative musical experiences. In *3rd Int. Conf. on New Interfaces for Musical Expression (NIME03)*, pages 129–134, May 2003.

[3] A. Cockburn and S. Jones. Four principles of groupware design. *Interacting with Computers*, 7(2):195–210, 1995.

[4] Mihaly Csikszentmihalyi. *Creativity : Flow and the Psychology of Discovery and Invention*. Perennial, June 1997.

[5] T. Djajadiningrat, B. Matthews, and M. Stienstra. Easy doesn't do it: skill and expression in tangible aesthetics. *Personal Ubiquitous Comput.*, 11(8):657–676, 2007.

[6] P. Dourish and V. Bellotti. Awareness and coordination in shared workspaces. In *CSCW '92: Proc. of the 1992 ACM conf. on CSCW*, pages 107–114, New York, NY, USA, 1992. ACM.

[7] H. Eden, E. Scharff, and E. Hornecker. Multilevel design and role play: experiences in assessing support for neighborhood participation in design. In *DIS '02: Proc. of the 4th conf. on Designing interactive systems*, pages 387–392, New York, NY, USA, 2002. ACM.

[8] M. R. Frieß, N. Klügel, and G. Groh. lzrdm: collaborative multi-touch sequencer. In *ACM Int. Conf. on Interactive Tabletops and Surfaces*, ITS '10, pages 303–303, New York, NY, USA, 2010. ACM.

[9] O. Hilliges, L. Terrenghi, S. Boring, D. Kim, H. Richter, and A. Butz. Designing for collaborative creative problem solving. In *C&C '07: Proc. of the 6th ACM SIGCHI conf. on Creativity & cognition*, pages 137–146, New York, NY, USA, 2007. ACM.

[10] E. Hornecker. A design theme for tangible interaction: embodied facilitation. In *ECSCW'05: Proc. of the 9th European Conf. on CSCW*, pages 23–43, New York, NY, USA, 2005. Springer-Verlag New York, Inc.

[11] T. Ingold. *Beyond Art and Technology: The Anthropology of Skill*. University of New Mexico Press, April 2001.

[12] S. Jordà and Marcos A. Mary had a little scoretable* or the reactable* goes melodic. In *6th Int. Conf. on New interfaces for musical expression (NIME06)*, pages 208–211, Paris, France, France, 2006. IRCAM — Centre Pompidou.

[13] S. Jordà, M. Kaltenbrunner, G. Geiger, and M. Alonso. The reactable: a tangible tabletop musical instrument and collaborative workbench. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches*, page 91, New York, NY, USA, 2006. ACM.

[14] M. R. Morris, A. Paepcke, T. Winograd, and J. Stamberger. Teamtag: exploring centralized versus replicated controls for co-located tabletop groupware. In *CHI '06: Proc. of the SIGCHI conf. on Human Factors in computing systems*, pages 1273–1282, New York, NY, USA, 2006. ACM.

[15] M. Puckette. Pure data: another integrated computer music environment. In *Proc. of Int. Computer Music Conference*, pages 37–41, 1996.

[16] J. Raskin. *The humane interface: new directions for designing interactive systems*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 2000.

[17] S. D. Scott, M. Sheelagh, T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *CSCW '04: Proc. of the 2004 ACM conf. on CSCW*, pages 294–303, New York, NY, USA, 2004. ACM.

[18] A. Tang, M. Tory, B. Po, P. Neumann, and S. Carpendale. Collaborative coupling over tabletop displays. In *CHI '06: Proc. of the SIGCHI conf. on Human Factors in Comp. Sys.*, pages 1181–1190, New York, NY, USA, 2006. ACM.