TreeQuencer: Collaborative Rhythm Sequencing - A Comparative Study

Niklas Klügel
Technische Universität München
Department of Informatics
Boltzmannstrasse 3
85748 Garching bei München
kluegel@in.tum.de

Gerhard Hagerer
Technische Universität München
Department of Informatics
Boltzmannstrasse 3
85748 Garching bei München
hagerer@in.tum.de

Georg Groh
Technische Universität München
Department of Informatics
Boltzmannstrasse 3
85748 Garching bei München
grohg@in.tum.de

ABSTRACT

In this contribution we will show three prototypical applications that allow users to collaboratively create rhythmic structures with successively more degrees of freedom to generate rhythmic complexity. By means of a user study we analyze the impact of this on the users' satisfaction and further compare it to data logged during the experiments that allow us to measure the rhythmic complexity created.

Keywords

Collaborative Music Making, Creativity Support, User Study

1. INTRODUCTION

In the last century the importance of rhythm has undergone a remarkable transformation from an equal to the most visceral element in the musical trinity of melody, rhythm and harmony - in both the high art and popular music [8]. Rhythm is social, it binds communities through dances they may share, such as in tribal or rave cultures, but also through participation in drum circles [10]. This participatory facet of rhythm, the collaborative creation of music, especially if intrinsically motivated, can be a highly engaging activity, *the joy of the groove* [12]. The related phenomenon of *Group Flow* [11, p. 158] is beneficial for social creativity, it is a motivator and means for the group to innovate in a creative task. In a musical context, it has been shown empirically to lead to more valuable results [6]. The personal experience of Flow itself, defined as "a holistic sensation that people feel when they act with total involvement" [1], is an enabler for the empathic involvement with music and stimulator for the implicit learning process [9].

We see these social factors as highly beneficial for musical creativity and therefore aim at supporting the collaborative creation of rhythmic structures while fostering Group Flow with help of a computer system. In such a setting, the computer can be seen as a mediator for the collaboration, helping to communicate musical intent or to facilitate the realization of musical ideas. Furthermore, in the context of group creativity, the heterogeneity of individuals, and their overall knowledge diversity, experience, and expertise are key elements [2]. The integration of these elements is also part of the mediating role of the computer. Focusing on the creative task itself, we aim at investigating the extent of rhythmic complexity and expressivity that users would want to pursue without being either underwhelmed or overwhelmed, as this is detrimental to the experience of Group Flow. Especially regarding the support for heterogeneous groups, this question is essential.

However, to our best knowledge, sufficient empirical results do not exist. Therefore we aim to perform a comparative study with several prototypical applications that successively give more degrees of freedom to the users with regard to shaping rhythmic structures and then evaluate their responses. We will furthermore measure the rhythmic complexity of the created rhythmic structures and compare them to these responses. In this regard, we try to shed some light into this issue such that later applications may be able to perform their role of the mediator more intelligently. We make use of a multi-touch table as shared interface, since this offers means that simplify social communication protocols [15, 4].

There have been several contributions that addressed the collaborative creation of rhythmic forms. One can distinguish between real time input driven systems such as BeatBugs [14] which allow users to directly perform percussive patterns that are then recorded into the system and sequencer based approaches such as the reacTable [5] or The Planets [7]. In case of the latter, users can modify the musical events asynchronously to the produced audio. We consider this temporal decoupling as beneficial since it conceptually allows for a larger variety of access points for collaboration and the mediating role of the IT-system. Therefore we design our prototypical applications as collaborative sequencing environments.

Additionally, we want to support melodic elements (e.g. bass sounds, simple harmonics) rather than primarily focusing on percussive ones, since we believe embedding the shared composition in a more musically comprehensive context reflects a more realistic use case. Furthermore, we seek to allow the modification of timbres in real time as timbre itself is an important part of the musical structure [3], especially for modern popular music (e.g. dance music).

2. CONCEPT

One of the most prevalent sequencing concepts in electronic music production is the step sequencer which dates back to the early 20th century (e.g. Raymond Scott's Circle Machine) and has been employed in various drum machines (e.g. Roland TR-808) but also in modern Digital Audio Workstations (e.g. Fruity Loops). In general, a musical duration is divided into a discrete set of units, the steps. At every step a musical event can be set to be triggered. This step sequence pattern is then played back in a loop. In many cases a 4/4 bar is divided into 8 or 16 steps, mean-

NIME'14, June 30 – July 03, 2014, Goldsmiths, University of London, UK.

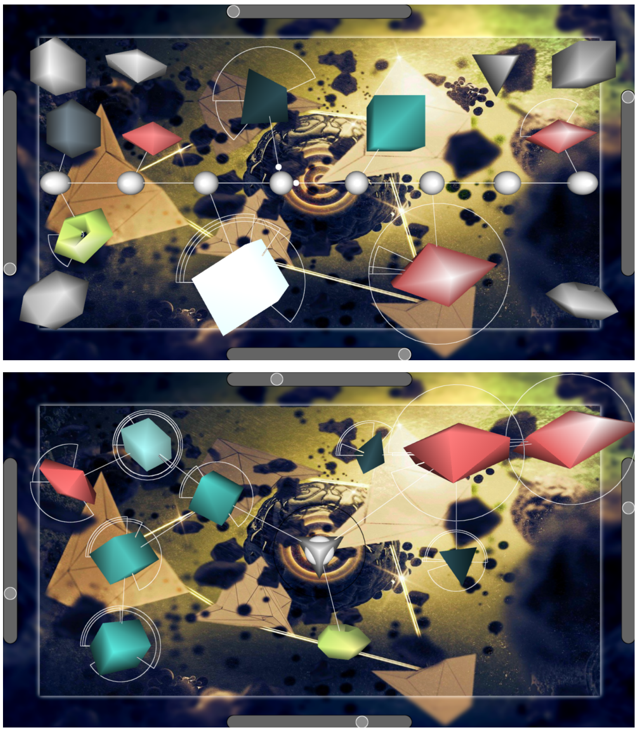

Figure 1: Screenshots of prototype 1 and 2

ing that events can happen at most at granularity of 1/8th or 1/16th. In this regard, a step sequencer is very limited in terms of rhythmic expressivity since only one time line exists that is fixed in length and because the musical subdivisions of a bar can only be even multiples of the duration of a step (e.g. triplets can not be used). We will make use of this core concept and successively extend it to allow for more degrees of freedom in regard to rhythmic expressivity for our prototypes.

The first extension allows to create polymeters such that looping rhythmic structures with even or odd overlapping elements can be constructed. This differs from the original concept as now several time lines (step sequencers) with the same pulse but different length are necessary.

The second extension concerns the duration of subdivisions in order to construct polyrhythms. This can be enabled by allowing each of the time lines in the polymeter to have a different pulse frequency. This makes it possible to create complex rhythmic patterns.

If we disregard the pulse frequency for a moment, then the step sequencer metaphor can be expressed as a graph construct in which nodes represent audible musical events (such as a hit of a bass drum) and edges express the succession of these events. In this way, a step sequencer with 8 steps can be seen a simple linear list of 8 nodes that are connected by edges. Similarily, a polymeter can be expressed as a tree that is to be traversed starting at the root. We will make use of this perspective on rhythmic constructs as interaction metaphor for our prototypes because in this way we can coherently express time-event relationships for all of our prototypes.

3. PROTOTYPES

In our prototypes, nodes represent sound generators directly, we chose to use different abstract three dimensional forms and colors to differentiate between them in the visualization (cp. fig. 1). These forms can be interacted with in the following way to change the graph structure and timbre:

- Dragging a form across the shared workspace (re-)connects it to the nearest form, thus this is used to construct the graph. Dragging forms outside of the workspace deletes them.
- A pinch gesture on the forms scales it and modifies its loudness. Forms can be muted when scaled to a minimum value, thus constructing rests.
- Rotation (2D or 3D using a multi-touch gesture) modifies synthesis parameters. Each form has 3 modifiable parameters associated. The global amount of a rotation of a form is indicated as one of three arc segments around it.

Edges of the graph are indicated as simple lines, the progression of the sequence(s) is indicated by small circles that travel along the edges from node to node at the pace of the pulse duration of the sequence. In this way it is visually indicated which forms are about to be triggered next. If this indicator reaches a form, the sound generator is triggered to produce an aural result. This is also shown visually by changing the brightness of the form according to the amplitude level of the generated sound.

We used six different sound generators that all allow to synthesize a large variety of different sounds. They can be controlled in real time therefore influencing the sonic output immediately. In this way users are able to alter the synthesized timbres and alter the musical result even when the rhythmic structure is static.

The first prototype p1, which is an adoption of a simple step sequencer, consists of 8 subdivisions represented as static, non-audible forms to which forms can be attached to. Only for this prototype no other graph constructions are allowed. Thus, only rhythmic constructions equivalent to that of a step sequencer can be manifested.

The second prototype p2 allows the construction of tree structures. The root of the tree is represented as non-audible form. Sub-trees leading from the root are looped according to the longest path from root to leaf node (each with the same pulse). This construction allows for parallel events to happen in parallel paths.

The third prototype p3 is a modification of the second one, where the pulse interval can be changed per sub-tree to a multiple of 1/8th. This is set by changing the distance of the first node of the sub-tree to the center of the workspace.

The global tempo in each prototype can be altered with one of four sliders placed around the workspace.

4. EVALUATION

For the evaluation we will make use of both a user study as well as tracking the use of the application by logging the interaction. With the latter it is possible to reconstruct the rhythmic patterns that have been created and therefore to apply measures for rhythmic complexity but also to analyze whether users alternatively preferred to shape the timbre of sounds via measuring the duration of gestures. 31 persons took part of our user study in groups of 3-4 people. With respect to their demography, their age was between 19 and 43 years, while 24 were male and 7 female. A large portion of our participants had a musical background (21 had performing experience). This is also reflected in the average time per week a test person makes music (9.4 hours). We performed the study by first presenting the second prototype and then the first followed by the third. Each of these sessions took 10 minutes after a short introduction about the basic usage. We tried to minimize any necessary instructions as we wanted to force participants to explore the prototypes. After the evaluation, participants were asked to fill out a questionnaire that contained questions regarding personal information and self-assessment, individual experience, group and interaction workflow as well as general feedback. Most questions used a 5-level Likert scale. Using the

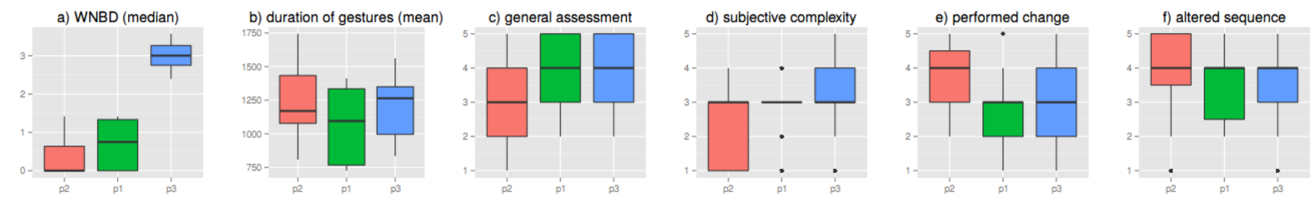

Figure 2: Plots of the logged data: a) Weighted Note To Beat Distance, b) the duration of gestures during interaction and the questionnaire: c) subjective assessment, d) perceived musical complexity, e) manner of changing objects (high means experimental, low means targetted) and f) the changed attributes (high means timbre, low means rhythm) for the respective prototypes p1 (green), p2 (red), p3 (blue)

logging data from the interaction with the applications, the graph structure at every time instant of the session can be reconstructed. This can be further transformed into a pattern notation in which every sound generator class is given its own staff and where all the trigger onsets of their respective multiple of a 1/8th are recorded. The length of the pattern is determined by the number of steps it takes until the whole rhythmic structure repeats (the smallest common multiple of the constituent patterns). After this transformation it is possible to apply rhythmic complexity measures for the whole pattern. We used the Weighted Note to Beat Distance measure since it was shown in [13] to agree with human perception and because it reflected more of the dynamic variation of the rhythmic structures in our data set compared to other measures.

5. RESULTS & CONCLUSION

In general, the feedback to all of our prototypes was very positive (cp. fig. 2c) with some of the groups playing up to one hour longer with the prototypes after the experiments. Several observations were made during the experiments themselves or when watching the corresponding recorded video material:

- Flow: Participants frequently moved to the beat by nodding head, clapping hands, tapping feet on the ground or the fingers on the edges of the tabletop. Even though people sometimes showed how much they enjoyed the rhythms by facial expression, many of these actions were done apparently unconsciously. The participants were especially engaged with the application when modifying timbres and discovering sounds (as they were shown only a small subset of the sound palette in the introduction). This was accompanied by showing emotions of amusement via verbal statements facial expressions (e.g. grinning). This shows that the exploration of sounds further caused constructive involvement leading to more engagement and curiosity, which are characteristics of a flow experience.
- Group Flow: Participants frequently showed not only positive emotions with regard to their own actions but also in regard to actions of their collaborators. We observed that the articulation of Flow of several participants at the same time led to a more extroverted expression of engagement such as vividly dancing to the produced music, collectively clapping to the beat or energetic laughing. In this sense it can be claimed that at times collaborators felt to be 'in the groove' with each other.
- Styles of interaction: In all groups a division of labor was observable. In many groups a single participant took the role of a coordinator, expressing his ideas while others helped executing them or assigning roles (as in a band setting) to his collaborators. This was done in an equally directed and undirected manner. The degree of verbal communication varied, some groups debated about the next steps while others did not talk at all but merely used gestures and eye-contact.
- Frequent Processes: We frequently observed the same sequence of steps that groups performed repeatedly to create their compositions. Namely *exploration*, where several persons experiment in parallel in a loosely coupled way. Followed by *restriction*, where after complaining about the non-musical quality of the aural output the group decides to remove all nodes and to restart the construction. *Planning* after a consulting conversation about improvement suggestions with the division of roles and working methods to avoid earlier mistakes. *Constructing* while the pace of the conversation fades and the interaction and collaboration becomes goal oriented. And finally *reception* of the accomplished if the musical result is felt worth listening while the group shows positive emotions of joy.

We now look at the data generated by logging the users' interaction with the prototypes. Regarding the question whether the additional degrees of freedom for rhythmical complexity have been used in a consequent way, the answer is negative on average (fig. 2a). The median values of the reconstructed rhythmical complexity values for each prototype show that the complexity generated by prototype two is less than for prototype one, although the opposite result would have been expected: for prototype one the maximum amount of rhythmic complexity that is producible is limited due to the fixed amount of "steps" and prototype two allows more the degrees of freedom. However, in accordance with expectation, the rhythmic complexity created with prototype three surpasses the results of the other two. These results agree with the results of the questionnaire elements asking for the perceived complexity of the created music per prototype (fig. 2c). One can conclude that the user's perception and the measurement of rhythmic complexity match, which not only indicates the validity of WNBD as an appropriate measure but also that the subjective user's assessment is also valid objectively. However, one question is whether the rhythmical complexities are voluntary or involuntary results. For this we will take a closer look at the results of the questionnaire.

The duration of gestures can be used as measure for whether users preferred to alter the timbre of sounds in real time rather than changing the rhythmic structure of the sound forms. Figure (fig. 2b) shows the mean of the duration of gestures for each prototype. We chose the mean here since it is susceptible for outliers and thus may better represent extended gesture input. Here it is apparent that in all applications timbral transformations of the sound forms were performed. However, regarding prototype one it is also visible that users may also have performed a large amount of rhythmical modifications.

For the evaluation of the questionnaire, we calculated Spearman's R as correlation coefficients for pais of answers. Regarding the collaboration itself, a high correlation is apparent between the ability to implement own ideas and fostering one's own creativity (0.71) as well as between the ability to implement own ideas and the preference to use the applications collaboratively (0.52). This means that users saw their collaborators and the applications as aids to pursue their own ideas. A highly positive correlation also exists between enjoying collaboration, on the one hand, and either evaluating the three apps as fostering creativity (0.55) or discovering new musical means (0.56) on the other hand. Users that expressed curiosity when working with the applications also expressed being inspired by the work of other users (0.61). Furthermore users who evaluated the musical outcome as harmonious or danceable liked not only the musical outcome in general (0.62, 0.52) but also liked the experimentation with sounds in a collaborative setting (0.36, 0.50). This may also lead to the conclusion that the collaboration may foster experimentation rather than the targeted construction of sound forms. This is apparent in the correlation between musical skills (a score accumulated from several biographical questions assigned to every user) and more critical evaluations of our prototypes: the higher the musical skill the more disappointment (0.37), confinement (0.45), and overstrain (0.46) for all applications were expressed, especially regarding the general assessment of prototype two (0.38). However, it seems unlikely that musically skilled users evaluated the applications negatively because of the actions of their collaborators, since the higher the musical skill the more positively collaborators and the interaction with them was seen (0.25). Many of our higher musically skilled participants attended a conservatory and thus may have substantial experience regarding musical collaborations. Given the correlation (0.32) that musically skilled users preferred to use the application three for a more targeted construction of sound forms and that users having stated that the applications exceeded their expectations especially assessed application one positively (0.65), it can be stated that for skilled users, application two is seen as too constrictive (corr. btw. skills and general assessment: -0.34) and application three as too difficult to comprehend (corr. btw. skills and comprehensibility: -0.23). The assessment of application one correlates positively with enjoying the experimentation with sounds (0.52). Additionally, there are positive correlations between the comprehensibility of the musical results generated with application one, on the one hand, and being able to implement own ideas (0.31) or experiencing pleasure (0.30) on the other hand. This comprehensibility of prototype one also correlates to perceiving the music as rhythmic (0.40) and to being inspired by other collaborators (0.39). This leads to the conclusion that application one in general is seen as more comprehensible and therefore traceable for collaboration. Additionally it seems suitable for sound-forms whose articulation is more rooted in timbre than in rhythmic (complexity) as can be found in today's dance music.

We assume that musically skilled users are adept to cope with and to create more rhythmically complex music. In this regard, a good prototype would allow such users to create rhythmically complex music compared to less skilled users. To gain more insight into such a relationship between musical skills and created rhythmical complexity, we performed a linear regression for these variables for each prototype. For prototype two, the slope of the best fit straight line is negative (coeff. -0.159), which may indicate that skilled users are not be able to express themselves accordingly while less skilled users may have created the complexity involuntarily. For prototype three the slope has no clear trend (coeff. 0.05), which may indicate that skilled users were unable to use their expertise for more targeted exploration. For prototype one, the slope has a stronger upwards trend (coeff. 0.24). Thus, from the viewpoint of the assumption previously made, this shows that this prototype is the one that scales best with users' musical skills. This more targeted interaction with prototype one, especially regarding the construction of the rhythmic structure, is also apparent in the users' evaluation of how they performed their changes (fig. 2e) and what was changed (fig. 2f). Furthermore this is also supported by the observation that the duration of gesture interaction includes a large portion of short events.

Concluding, it is to say that it is important to gather more empirical data in order to be able to make further reliable statements investigating the appropriate level of expressivity for collaborative composition support. This is our intention for future work.

6. REFERENCES

[1] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper and Row, New York, NY, USA, 1990.

[2] G. Fischer. Distances and diversity: sources for social creativity. In *Proc. 5th Conf. Creat. Cogn.*, pages 128–136. ACM, 2005.

[3] P. Holmes. An exploration of musical communication through expressive use of timbre: The performer's perspective. *Psychol. Music*, Mar. 2011.

[4] E. Hornecker, P. Marshall, and Y. Rogers. From entry to access: how shareability comes about. In *Proc. conf. on Designing pleasurable products and interfaces*, pages 328–342. ACM, 2007.

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

[6] R. MacDonald. Creativity and flow in musical composition: an empirical investigation. *Psychol. Music*, 34(3):292–306, July 2006.

[7] M. Meier and M. Schranner. The Planets. (NIME 2010):501–504, 2010.

[8] B. Neill. Pleasure beats: rhythm and the aesthetics of current electronic music. *Leonardo Music J.*, 12:3–6, 2002.

[9] L. Nijs. *The musical instrument of natural extension of the musician*. Witwatersrand University Press, 2009.

[10] J. Pareles. The Rhythm Century: The Unstoppable Beat. *New York Times*, pages 2–3, 1998.

[11] R. K. Sawyer. Group creativity: musical performance and collaboration. *Psychol. Music*, 34(2):148–165, Apr. 2006.

[12] B. Swift. A Chasing After the Wind : Experience in Computer-supported Group Musicmaking. *OHZCHI*, pages 1–6, 2009.

[13] E. Thul. *Measuring the Complexity of Musical Rhythm*. PhD thesis, McGill University, 2008.

[14] G. Weinberg, R. Aimi, and K. Jennings. The Beatbug Network: A Rhythmisc System for Interdependent Group Collaboration. *Proc. NIME 2002*, pages 107–111, 2002.

[15] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann, 2011.