Cross-Modal Parametric Composition*

David Gerhard and Daryl H. Hepting
Department of Computer Science, University of Regina
{gerhard, dhh}@cs.uregina.ca

Abstract
We propose an interactive composition technique to facilitate the exploration of complex parameter spaces. More computer-aided composition than algorithmic composition, human aesthetic judgment is an integral part of the technique. From a set of central parameters, the composer chooses a mapping which allows the generation of both audio and video representations. Features from each representation are then extracted and used as parameters in an iterative re-render process, each step of which can be automated or respond to human input. With this technique, composers can see as well as hear their music as it develops, and can interact and follow paths through the parameter space, exploring sound spaces heretofore unavailable to standard compositional methods.

1 Introduction

Computer related composition has historically fallen into two broad categories: computer aided composition, and algorithmic or automatic composition. In the latter, the algorithm is set free to compose new works independent of a human, and the final work is then judged for aesthetic appeal. Computer aided composition, in contrast, provides algorithmic tools to composers, allowing them to interactively build new works based on personal, musical or intuitive factors. The problem with this type of approach is that the composer has a great deal of responsibility for choosing which tools or techniques will be used and how they will be invoked, and this type of decision making process takes a great deal of patience, since multiple (often similar) alternatives must be considered before a decision can be made.

This work focuses on computer aided composition, and discusses a new form of composition tool which allows the composer to explore musical space from a variety of modes and perspectives and to choose between a set of like options, thereby maintaining the human facility for aesthetic evaluation within the composition development cycle. Since the options to be selected by the composer are automatically generated, this method is in part automatic composition as well as computer aided composition.

We discuss the composition process as an exploration of parameter space, with each set of musical parameters corresponding to a point in that space. The parameters are initially decontextualized, meaning that they can represent any musical or aesthetic quality or process, to the point that the mapping and re-mapping of these parameters to meaningful musical constructs is very much a part of the composition process.

This technique an improvement upon historical computer aided composition techniques because it regularizes the exploration of these parameter spaces, allowing the composer to navigate and explore in an efficient manner without having to consider every possibility in the space. Similar to the concept of a non-exhaustive search, the idea is not to present every possibility to the composer, which would be overwhelming and eventually useless. Instead, the idea is to present a set of increasingly specific representative possibilities to the composer, and as selections are made, the composer “homes in” on an interesting area of the parameter space and can explore multiple resolutions.

The proposed technique contains the following components which allow for improved access to previously unavailable parameter space explorations:

1. The cogito interface (described below)
2. Cross-modal parameter exploration
3. Iterative secondary parameter extraction and mapping

In a way, the cogito system and the accompanying parameter exploration technique is middleware, in that it serves as an intermediary between a composer and a set of modern composition assistance techniques such as automatic counterpoint generation. Rather than having to deal with the parameters associated with the invocations of those algorithms, the composer can deal with them at the conceptual level and evaluate the intermediate results of each technique.

Like other fields, the technician’s view (the developer of the algorithm) is different from the user’s view (the com-
poser), and this gap often detracts from the direct usability of the composition tool, and hence from the composition activity itself. Rather than having the system free the composer to compose, the composer is forced to deal with the minutia of the particular implementation. In contrast, the proposed technique would allow the composer to work within his or her own conceptual framework for the composition, and thereby work directly with the music.

2 Parameterization

Many composers have explored the connection between modes of interaction with artistic pieces, and the recent increase in artistic multimedia installations shows that interaction with multiple media composed on the same topic adds to the artistic experience. With this motivation, we propose the concept of cross-modal composition, where a composer who works primarily in one field (a musician) may benefit from exploring the musical parameters of the composition in another context. In this case, and in the proof-of-concept we discuss below, we use the visual context as an augmentative media to a musical composition.

The composer builds a set of paths through parameter space, and these paths are used to generate not only music, but also a corresponding animation which the composer then uses to aid in the further exploration of the parameter space. The musician need not know anything about the development of visual animations—the system would be able to create novel animations based on the parameters chosen. A further benefit of this system is that musicians and animators can use the system to collaborate toward a cohesive multimedia work which is rooted neither in sound nor in vision, but in the central parameters used to generate both media.

This is not generative music per se, since the composer maintains control of the parameters used to generate the music and visuals, as well as the mappings used to move from the parameters to the musical or visual space. The computer system then finds similar areas in the parameter space, which are expected to generate aesthetically similar musical constructs.

By allowing the composition to occur in different modalities, or across modalities, we give the composer an opportunity for new perspectives. The composer may see something in the visual representation of the piece that triggers a new idea to be added to the acoustic representation. This can be related to the oblique strategies concept (Eno and Schmidt 1975), where a composer or artist who might be stuck with so-called writer’s block can choose a strategy from a deck of cards, and this will suggest an action to explore, such as “Fill every beat with something” or “allow an abandonment of structure”. These strategies allow the artist to pursue lateral thinking and to come up with new ideas. The oblique strategies concept is a general form of lateral thinking, while our proposed cross-modal composition tool is much more specific to the piece of music being composed. Because it takes into account the central parameters being used to generate the work in the first place, cross-modal composition allows the composer a different way of looking at the parameters and the composition as a whole.

Parameterization allows a composer to work in the abstract without needing to realize the ideas or composition in any specific way. Composers such as Dahlstedt (2001) have considered composition in parameter space. Roads (1996) offers many examples of parameterized computer-assisted and automatic composition. Parameter-based composition allows the composer to navigate a large space of musical ideas and pick what is aesthetically desired. Stockhausen (1989) explored different parameters in his music, and Boden (1996) discusses conceptual spaces defined by stylistic conventions, and how to expand those spaces with three rules: “drop a constraint”, “consider the negative”, and “vary the variable.”

2.1 The Composer and the Programmer

The abstraction to parameter space can be compared to the building of a computer program. The program must eventually exist as machine code, otherwise the computer will not be able to perform the events specified in the program. Despite this, the programmer must interact with the program at a much higher level, since most humans are not efficient processors of machine language. Current programming techniques require a programmer to write commands in a high-level programming language, which is the same idea as writing notes in a musical score - the orchestra then converts those notes to specific actions on their specific instruments with the goal of making specific sounds. Similarly, the computer converts the higher-level language code into specific actions to be performed on specific pieces of hardware - read from a disk, display on a screen, add some numbers together, and so forth.

At the highest level is the algorithm on which the program is based, which is independent of the programming language, the computing platform, or in fact of the concept of being computed. The algorithm could be executed by hand or with water-wheels or with anything that can accomplish a task. In the same way, the parameter space is completely removed from and independent of any realization in musical space. This analogy is summarized in Figure 1.

Visual programming is a newer technique where the programmer moves boxes around and connects them with lines, producing data paths. Each box is a specific high-level operation which the computer will turn into low-level commands to specific hardware devices. The algorithm is then represented
3 Parameter Space Exploration

When composing a piece using this technique, the composer begins by assigning a set of central parameters. These central parameters can be specified in any way, and this allows greater flexibility in the composition techniques. A composer may draw a few central ideas on a computer screen or tablet, or move an input glove in a virtual reality environment, or may play a few notes on the instrument with which they are most familiar, or perhaps submit a few mathematical equations or algorithmic ideas. The versatility of this method is that the central parameters are a complete abstraction from the input technique, and as such any input technique can be used to build a set of central parameters.

Once the central parameters are in the system, the musical interpretation of these parameters is again entirely flexible—parameters can indicate note events, acoustic properties, in fact any sound based idea, and it is up to the composer to choose (with the aid of the system) a mapping from parameters to acoustic properties which will result in a cohesive work conveying the composer’s intentions. Parameters can be duplicated and re-mapped to different modalities to generate cross modal fugue-like structures—a parameter curve can represent a theme which is realized in a pitch-based mode, and then re-appears later in the piece in a timbre-based mode or a rhythm-based mode, so that the same theme is repeated in different parameter mappings. These mappings can be investigated and explored as a unit or individually, so that if a particular theme needs to be revised slightly for it to “work” in a different parameter mapping, this is possible through the iterative refinement of the cogito system.

There is practically an infinite variety of visual representations for any set of data, and indeed the dimensionality of the parameter space proposed with this technique is limited only by the composer’s parameter choices and mapping schemes. Bertin (1983) wrote that “to construct 100 DIFFERENT FIGURES [his emphasis] from the same information requires less imagination than patience. However, certain choices become compelling due to their greater ‘efficiency.’ ”

Bertin (1983) and Larkin and Simon (1987) define efficiency in terms of the effort required to comprehend an image. Following the distinctions of Ware (1993), there are both sensory and conventional aspects of efficiency: sensory efficiency is maintained if perceptual cues are used in standard ways, but conventional efficiency is maintained only if the viewer is practiced in interpreting the particular convention. Efficiency may be reduced if sensory cues are used improperly. However, there can be much to gain if conventions are challenged.

Since all aspects of the composition process can be selected and explored using this paradigm, the cross-modal interaction can also be explored and evaluated. If there was an
apparent dissonance between the visuals and the sound, the composer could re-examine the mapping to each mode, but if the acoustic composition is approaching what the composer has in mind, then on some level the corresponding visuals are less important—even if the animation is not “efficient” (in the sense described above) and is hard to comprehend but still leads to a valuable composition, then the idea of efficiency in this context is not as important. The cohesiveness and efficiency of interacting with the piece is important only in the context of the modality of final presentation. Efficiency in musical terms has not been investigated in as much detail, but in the final evaluation, efficiency cannot be removed from the concept of task, or from the audience viewing or listening to the work.

An advantage, therefore, of exploring a musical space using a computer-aided parameter exploration technique such as this is that the space can be explored more fully and with less effort than with manual experimentation. It is far easier to have a system that will generate a set of variations for the composer, and the composer has merely to choose between them. The parameter space is navigated by example, and many variations can be investigated which might not otherwise be available to the composer.

3.1 Cogito

Cogito (Hepting 2003; Hepting, Fracchia, Dill, and Russell 1996) has been developed as a means to preserve the user’s ability to exert judgment while exploring a large, complex parameter space. Points in the space are constructed from all possible combinations of values from the different parameters. For example, given 3 parameters each with 10 possible values, the space has $10^3 = 1000$ points. Foley and Ribarsky (Foley and Ribarsky 1994) wrote that only automatic methods could be used effectively with large parameter spaces because otherwise the users would be overwhelmed. Users of the cogito system have not been overwhelmed, however, even though they deal with millions of possible combinations (Hepting 2003). The system allows users to apply their own judgment when assessing alternatives automatically generated by the system, which has the patience to realize these different combinations and keep track of them.

The system functions in two ways: first it allows people to explore different relationships among parameters. By reorganizing the view of the parameter space (as shown in Figure 3, it is possible to transform one’s perception of the parameter space from Klondike (difficult to navigate) to homing (easy to navigate and find a desired solution) (Perkins 1995) Secondly, it records and manages the exploration of space, so that the space can be explored in a systematic way and one can return to previous explorations to examine other options at decision points. In this way, the system is an example of a numerical experimentation environment (Ioannidis, Livny, Gupta, and Ponnekanti 1996). Where there are examples of interest, the user may click to select them. The subsequent space of alternatives contains all possibilities consistent with those selections. For each parameter, the list of acceptable values comes from those in the selected exemplars.

The Design Galleries concept (Marks et al. 1997) is related to our cross-modal composition technique. In Design Galleries, a user builds an animation or a figure, then wants to choose a lighting condition or an object position or some variation on one or two parameters. The system offers a gallery of choices that represent the range of possibilities on that parameter, and the user selects one that is appropriate for their purposes. This concept could be extended to musical composition, where a composer wishes to vary a musical parameter relating to a piece being composed.

The difficulty with Design Galleries is that it requires a great deal of information to be available before the system is used - the range of parameter values must be defined, and the majority of the composition must be completed if the parameter variation is to be seen in any sort of context. Further, if the composer discovers that it is the parameter mapping itself that should be modified, they must go to the beginning of the Design Galleries process and re-define the parameter mapping.

Design Galleries can show alternatives that are good examples of a set of values, however the evaluation function has to be specified a priori. The cogito system permits results to be partitioned based on qualitative or quantitative differences in values within one or more parameters.

![Figure 2: Schematic look at the hierarchical interface: the space of available alternatives is grouped according to user-specified criteria. Each group (A – F) has a representative element (a – f) which is displayed to the user. The subspace for the next search iteration is based on the user selection (b and f).](image-url)
Figure 3: Consider a three-dimensional space, depicted in the top left, with axes X, Y, and Z. Organizing the space in terms of any of those 3 axes leads to the other states depicted. If elements in component X are chosen sequentially, those in Y and Z can be selected randomly to give a sense of the available options.

3.2 Keystate Path Generation

A further augmentation allowed by this process is keystate path generation. With the aid of the computer, the composer finds some areas in parameter space which would be interesting to explore. Each individual parameter space location is quantified by a keystate, similar to the idea of keyframes in animation. The composer can then indicate a pair of keystates and the system will compose a number of potential paths between the keystates, which can be explored by the composer. Each path between keystates represents a musical theme or phrase, and these paths are combined to create the finished work. The paths between the keystates in effect become another parameter in the composition, and the composer can choose between a set of paths and iteratively narrow the path space in the same way that the parameter space mappings and the parameters themselves can be explored. The keystate path exploration paradigm is shown in Figure 4.

In traditional western musical composition, a composer may want to modulate from one key to another, and there are some standard ways to do this, some standard chord progressions for example, which are learned through experience or formal training. When composing in non-traditional sound parameter spaces, it may not be immediately obvious which transitions are available between one key state and the next, or between one general area of the parameter space and the next. The system then aids the composer in exploring the available transitions, and in choosing one transition path over another for various artistic and aesthetic reasons. The computer then does not make any artistic judgments, it provides a concise and efficient way to explore the parameter space, which otherwise would have to be explored state-wise by the composer. In addition, this method ensures that the composer will not miss any potentially relevant state or state transition, because as the system presents a representative subset and the composer chooses states or transitions that are closer and closer to what the composer might have in mind, the relevant parameter space is explored and the irrelevant parameter space is ignored.

This technique is not unlike the concept of computational steering, where a computer algorithm is executed in real time and a visualization is made available to a user or programmer. The user can then track the progress of the algorithm and make decisions about parameters. If the algorithm does not appear to be converging, the user can alter parameters during the computation, which may lead to a successful result. In this way, the proposed compositional system could easily be modified to create an interactive performance system as well, where the user initially creates a set of parameters (like a score) and then interacts with them in a real-time context to produce a live performance work.

3.3 Computational Path Composition

The potential paths from keystate to keystate can be generated using any of a number of current automated composition techniques. Many of these are discussed in Roads (1996), and some will be described here. The choice of composition technique or a combination of various techniques is also available to the composer, and different techniques are better suited to different parameter spaces and mapping paradigms. Algorithmic composition systems can be categorized in several ways. Stochastic composition tools incorporate probabilities and random numbers in the generation of acoustic
events, while deterministic systems use rules and procedures such that the same input will always produce the same result. Researchers have experimented with fractal composition and noise shaping based on human perceptual properties.

Birkhoff’s aesthetic theory (Schroeder 1991) states that for a piece of art to be interesting, it should be neither too regular nor too surprising. Many composers have recognized that in artistic pieces which meet this criteria, parameter distributions corresponding to \( \frac{1}{2} \) are often discovered, meaning, for example, that the frequency of the notes in a piece (or often of the intervals between the notes) is related logarithmically to the past frequencies (or frequency intervals). This is often called a “pink” process because it has the same distribution as “pink” noise, which is between “white” noise, where each event is completely independent of all past events, and “brown” noise, where each event is linearly dependent on all past events. Many other compositional techniques are currently available, including genetic algorithms, cellular automata, Markov processes, neural nets, and emergent behaviors. The advantage of this parametric composition technique is that the rendering of paths between keystates can be done with any or all of these techniques, and the choice of these techniques or combinations thereof is limited only by which techniques are implemented and available in a specific technique being used by a specific composer. If a composer is interested in investigating the parameter space available with genetic algorithms, for example, this method could be used for all path interpolations and other parameters could be varied. If the composer is exploring the locations around the keystates themselves and is not as deeply interested in the transitions between them, then the assistive composition technique is less important.

### 3.4 Secondary Parameter Extraction

Once the central parameters are defined and a mapping is found, the system renders the various media, as defined by the composer. A composer may choose to render audio and video, or audio and text, or dual streams of audio with different mappings from the central parameter set. Whatever the choice, at this stage all media are generated from the central parameters and the parameter mapping. Once the media has been generated, and the piece is available as a whole, the composer can interact with the various parameters, and can extract parameters from the various media. If, for example, the alternative visual modality corresponding to the composition has provided some interesting effects, the composer can extract parameters based on these effects from this secondary modality and use them as parameters in the re-generation of the composition. This procedure is fully iterative and can continue indefinitely, as refinement or as space exploration.

The procedure is presented in Figure 5.

An example of the result of such an extraction revision is presented in Figure 6. In this example, a set of central parameters was defined as input to an animation generation engine, and the musical composition was generated as a secondary modality to the animation. The animation was based on a set of fractal transitions from a standard form, a Sierpiński gasket. Each transition was based on the set of central parameters. The result was a single figure that moves and changes shape across time. The music was generated from a set of rhythmic loops, based on the same central parameters. Information was then extracted from both the resulting animation and the resulting soundtrack, and used as parameters to regenerate the animation and the soundtrack. We call the set of parameters in this format a score, and note here that this score is interactively created by the repeated exploration of keystates and transitions between them, of parameter mappings to realized media, and of extraction of secondary parameters to be re-applied to the media generation.

![Figure 5: The process of generating media from a parameter space. Begin by choosing a central parameter variation which can be mapped separately to parameters in the various modalities. Once the media has been generated, intermediate parameters can be extracted and mapped, and the central parameter set can be re-mapped until a desired effect is achieved. At each iteration, the generated media is available.

The process we followed to generate our proof of concept is described here:

0. Choose a theme for composition. We chose the rotation angle of fractal figures

1. Explore different images relating to this central theme. We examined various rotational parameters relating to these still images
Figure 6: Score of parameters. $R(T_n)$ is the rotation factor for each transformation. $F$ scale is the scale factor for the entire figure. $T$ scale is the scale factor for the transformations. Amplitude is the power level in each frame of the generated sound file, an extracted parameter. Size is the number of non-black pixels in each frame of the generated animation, an extracted parameter.

2. Construct a central parameter variation for the rotational parameters. This is shown in the top three frames of Figure 6

3. Render a complete set of images and generate a movie

4. Tweak various parameters including starting angle and rotational speed

5. Repeat steps 3 and 4

6. Generate music from the central parameter variation

7. Extract parameters from the images and the music. The result of the parameter extraction is shown in the bottom two frames of Figure 6

8. Add variations and decorations. Shapes that we identified for decoration addition are shown in Figure 7

The final composition, entitled “Triangularhythmic” including both video and audio components is available for viewing in the ICMC 2004 Digital Jukebox interactive installation.

4 Discussion: The Grand Structure

In composing works using parameter space exploration, it is potentially easy to get bogged down in the details and not see the entire work. For this reason, it is important to occasionally step back and view the parameter space as a whole, with keystates and paths snaking through and around in an intricate picture. It is possible then to take the whole system and change some overlying parameter, such as “Make the entire piece in 5/4 time” just to see what might happen. At this higher resolution, it is also possible to investigate multiresolution paths and parameters - the entire piece might concentrate around three main keystates, moving from one to another, or orbiting and exploring a particular area of parameter space. Within the higher resolution of the area around that keystate, there might be several keystates and paths between them. A composer can explore higher-dimensional parameter spaces in this context by investigating a lower-dimensional projection of the space.

It is also important to track a composer’s actions historically over time, so the composer can return and re-visit any decision or specification that was made in the past. The overarching history of changes can be explored itself as a parameter, or used to see the progression of a work from the general to the specific, or through variations or interpretations of a theme.

Ultimately, the utility of any computer-aided composition technique will be judged by the quality of music that composers can create by using it, and by the ease with which composers can create music with it. The versatility of this technique is also its customizability - if a composer uses this system, becomes familiar with it and builds into it a set of commonly used techniques and themes, it becomes an extension of the composer and merely a tool to augment the abilities of the composer, not to replace them.
5 Acknowledgements

The authors wish to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada, as well as the University of Regina. Thanks also go to Matthew McKague, who did much of the programming for the animation proof-of-concept.

References


