Bridging Design Thinking and Creation

Gestalt Editor: A constructive software environment

Takayuki Hamano*, Kiyoshi Furukawa**, Haruyuki Fujii***, Kazuo Okanoya****

** JST, ERATO, Okanoya Emotional Information Project, hamano@japan.com*
*** Tokyo University of the Arts, kf@zkm.de*
**** Tokyo Institute of Technology, fujii.h.aa@m.titech.ac.jp*
***** JST, ERATO, Okanoya Emotional Information Project, kazuookanoya@gmail.com*

Abstract: Many artifacts created by humans presumably embody two kinds of structures: the surface structure and the deep structure. The former includes the composition of the created artifact and the direct operations for that composition, while the latter consists of three levels: (1) the level of the abstract structure of the composition process of the expression experienced, (2) that of the abstract structure of experience, and (3) that of emotions arising from experience. The deep structure is assumed to be related strongly with human cognitive processes. Our hypothesis is that different manners of expression may have commonalities and differences in the deep structure.

As a first step of elucidating these structures, we developed *Gestalt Editor*, an experimental creative software environment that connects the surface and the deep structures. Gestalt Editor was developed both for creation and research, facilitating the design and the exploration of the deep structure and bridging it with a pluralistic artifact. In order to verify the commonalities between different expressions, Gestalt Editor allows a user to explore music and space representations and generate both experiences based on the common deep structure.

Key words: constructive approach, surface and deep structure, music and space generation

1. Introduction

In general, the creation of art is realized as a pluralistic expression by composing various elements such as motifs and materials. That compositional process is mostly dependent on a creator’s experience and skill, i.e., tacit knowledge. Many have tried to elucidate the secret of creative processes, but it has never been explained clearly. We attempt to establish a methodology to discover a commonality in creative processes based on a constructive approach.

1.1 Surface structure and deep structure

There are different layers of structure in a composition: the surface structure and the deep structure (Figure 1). The surface structure is directly related with the expression itself and its components including concrete elements such as timing, pitch, and duration in music, or shape, scale, and size in space design. Another layer is the deep structure, which is associated with forms of an artifact, and it consists of three levels: (1) the level of the abstract structure of the composition process of the expression experienced, (2) that of the abstract structure of experience, and (3) that of emotion arising from experience.

The authors were inspired by Chomsky’s government and binding theory for constructing the notion of deep structure and surface structure. In the government and binding theory, the D-structure encodes the basic theoretical relations and lexical properties in a sentence as determined by the argument structure of the predicate, and the S-structure is closer to the surface manifestation of the sentence [3]. We believe that this theory can also be applied to music and space design with regards to the fact that the surface structure deals with semantics whereas the deep structure deals with the syntax of an artifact.

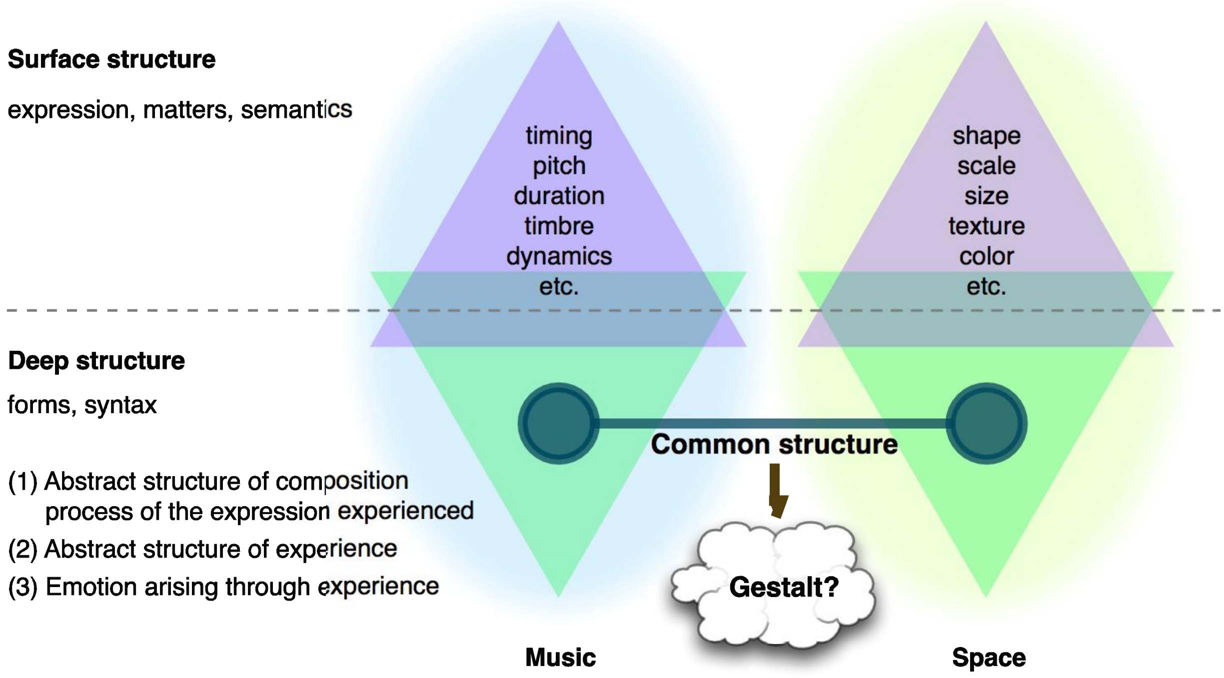


Figure.1 Surface and deep structures in a composition

1.2 Commonality in deep structure

We assume that diverse kinds of expressed artifacts, such as music, verbal expressions, and spatial organization such as buildings and gardens, have commonalities and differences in their deep compositional structure. The deep structure is considered to be related with human cognition in the process of designing the expression. In addition, the deep structure possibly affects implicitly the cognitive process of understanding expressed works. If there is an underlying common compositional principle, it might be something that forms a gestalt in the mind and produces cognitive sense perceived as more than the sum of the parts of the composition.

Our hypothesis above is based on a previous work by Furukawa, Fujii, and Kiyomizu [2]. Their work proposes that there is a particular relationship between the manner in which we experience music and the way in which we experience spatial structure, e.g., the traditional Japanese garden and Japanese architecture. This relationship encompasses both similarities and differences. Their study shows that a comparison of time-based flow of music with passage along a garden circuit provides us with material from which we can distil several unique perspectives. On the basis of these perspectives, we constructed our system to determine the foundation common to both musical and spatial experiences.

1.3 Related works

There have been some theoretical studies on expression and its composition; for instance, in the domain of music, the *Generative Theory of Tonal Music* proposed by Lerdahl and Jackendoff [4], and in visual design, *Shape Grammar* by Stiny and Gips [8] are well-known theories. Both of them deal with the surface structure, i.e., the syntax of the expression.

There are also the other approaches that associate music with spatial organization. Noted works are *Furniture Music*, coined by Erik Satie, which makes music a part of daily life [7]; *Soundscape* [6] by Murray Schafer, which attempts to integrate the visual and sound environment; and asynchronous mapping of sound into architectural structure by Le Corbusier, Varèse, and Xenakis [11]. All of these works have taken a practical approach to create new expressions.

In recent studies, Fujii, Furukawa, and Kiyomizu [1] explored the similarities and differences between the structure of operations in architectural design and that of operations in music composition. Their paper describes an experimental system to grasp the relations between music and spatial design. Their system observes the sequence of scenes in an abstract space symbolically representing an architectural space and translates the experience into the experience of the structure of music.

Another example is a space–sound inter-media system as a tool for encouraging meta-cognition and inspiring embodied experience in space by Suwa and Fujii [9, 10]. This system, based on meta-cognitive verbal reports on spatial experience, assigns sound sources to objects constituting a space and generates a piece of music using rules for mapping verbal structures onto sound structures. They hypothesized that having people listen to the generated piece of music on the spot may encourage meta-cognition and facilitate re-experiencing the same space in a different manner. In order to explore the feasibility of the hypothesis, they conducted a cognitive experiment with the authors themselves as participants.

1.4 Goal of the project

The ultimate goal of our project is to explore the commonality of compositional principles in a variety of representations, particularly music and space organization. This will be achieved by investigating the relationship between the surface structure and deep structure. Therefore, we would like to shed light not only on the actual creative process, but also on the process of design thinking that occurs before creation.

2. Method and Expected Realizations

2.1 Method

The general processes of scientific research capture the characteristics of target objects and phenomenon analytically, form a hypothesis regarding general rules explaining that phenomenon, and verify its validity by objective observation. However, we believe this constructive approach would be more effective for research as in our project. The main reason is that artistic expression often involves subjectivity and diversity, and it should be helpful to carry out observations through the accumulation of experience rather than under limiting conditions.

Therefore, we planned our research procedure as follows. First, we developed an experimental software environment that visually shows the compositional structure in a logically calculable way, and generate music and images of space organization automatically. Then, we asked many musicians and space designers to create works using our system in order to collect data about the composed structures. We analyzed the collected data and tried to discover the commonality of the structures. Those analyzed results may be returned to the creators as part of a new idea.

2.2 Expected results

The results of this study will elucidate not only the actual composition of the completed music and space, but also the commonalities and differences in the structure of the design techniques used in creative processes. The study will also contribute to devising a new expression that connects different manners of expression taking cognitive aspects into consideration, as well as designing the expected experience. Furthermore, it will facilitate skill acquisition by making the compositional processes explicit rather than intuitive as it is with professional artists and designers.

3. Implementation of software environment

We developed a software environment, *Gestalt Editor*, which enables the design and exploration of the deep structure, bridging it with a pluralistic artifact. Users can explore music and space representations and generate both experiences based on the common deep structure. This environment is newly developed based on the idea of a previous version [5].

3.1 Process of creation

First, we designed a procedure model of the creative process flow using Gestalt Editor as shown in Figure 2.

Gestalt Editor is able to simulate all of the following processes. The basic idea is that the user builds a composition by combining motifs or materials. The content of those elements can be modified by operation. Abstraction is an optional function to create a more complex composition. When the user tells the environment to generate a composition, it is produced as an artifact. The system is also capable of describing the process of recognizing that artifact and the experience caused by that recognition.

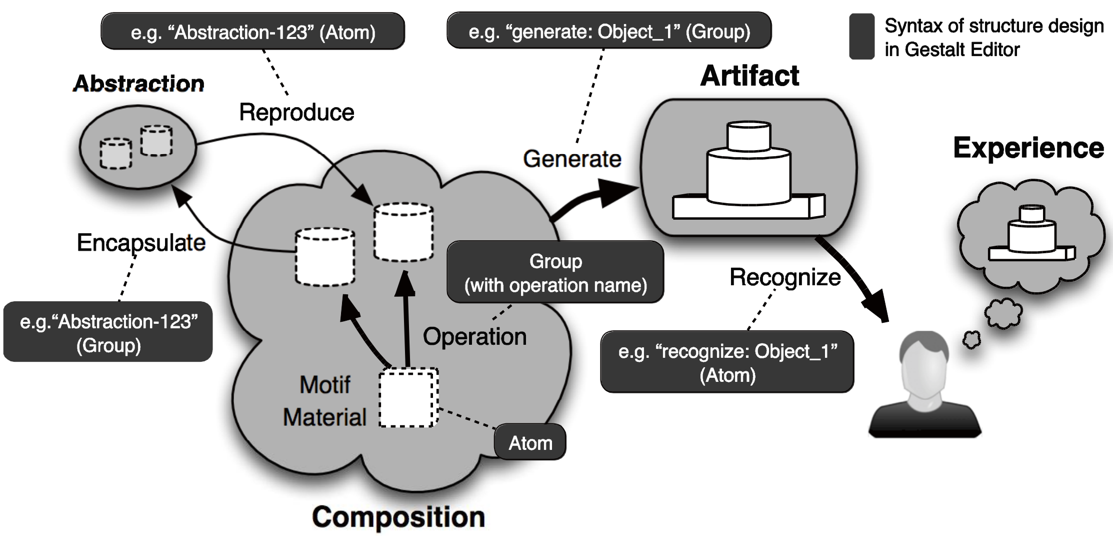


Figure.2 General flow of the creation processes using Gestalt Editor

3.2 Structure composition

In Gestalt Editor, the user builds a structure using the program called *Editor* (Figure 3). The editor has a two-dimensional canvas. The user is allowed to place two kinds of visual elements on the canvas, *Atom* and *Group*.

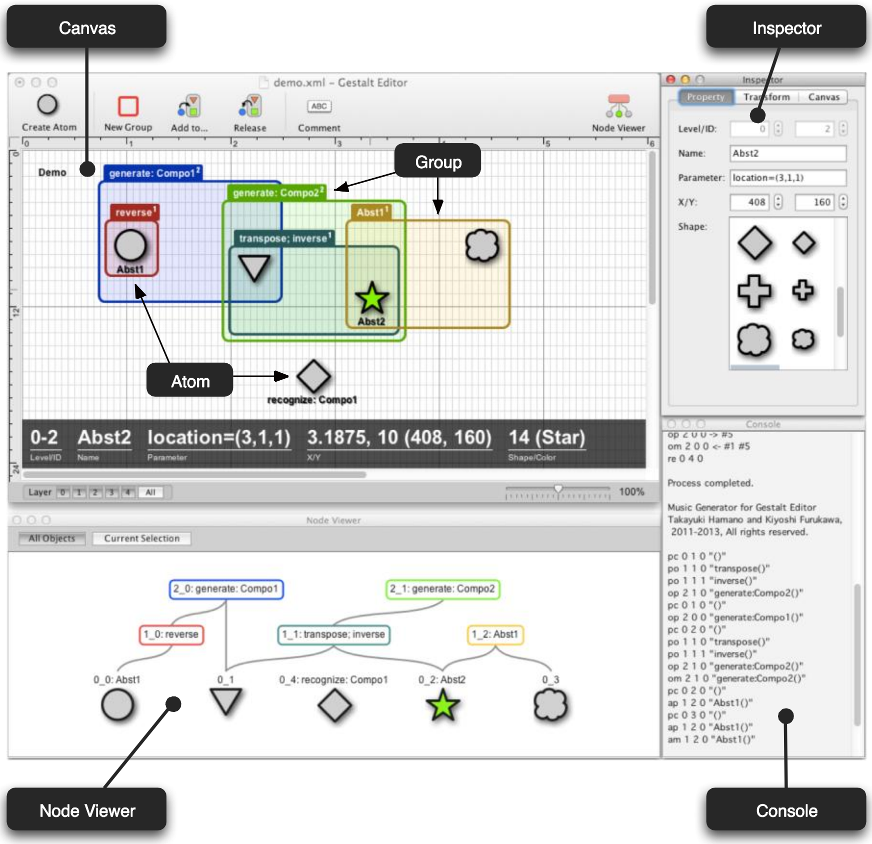


Figure.3 Screenshot of Gestalt Editor

3.2.1 Basic compositional elements

Atom is the most primitive compositional element shown on the canvas and is shown as a simple shape such as a circle, triangle, or star. The most basic usage of Atom is to represent the creation of a motif or a material of the artifact. The shape of Atom represents the kind of motif to be created.

Group is a function to create a group of elements. It is also possible to create a new group containing other groups. It is used to declare common attributes between elements, as well as to modify the content of motifs created by Atoms. This modification of motifs is called an *Operation*.

Using these elements, the user can design a structure representing the relationships between each pair of elements. When building a complex structure, *Node Viewer* allows the user to check the relationships between elements in detail.

3.2.2 Attributes

Both Atom and Group have some attributes such as *level*, *id*, *name*, *parameter*, and *location (x, y)* (Table 1). These attributes are used in many ways in the later generation processes.

Table 1. List of attributes of Atom and Group

Attributes	Description
level	depth of structure created by groups
id	identification index number
name	name of element
parameter	optional parameters
x	horizontal location on canvas
y	vertical location on canvas
shape	(only for Atom) shape of Atom
color	(only for Group) color of Group

3.2.3 Advanced notations

There are some special functions for composing a structure using Gestalt Editor. These functions are identified by writing special notations on the names of elements (Figure 4).

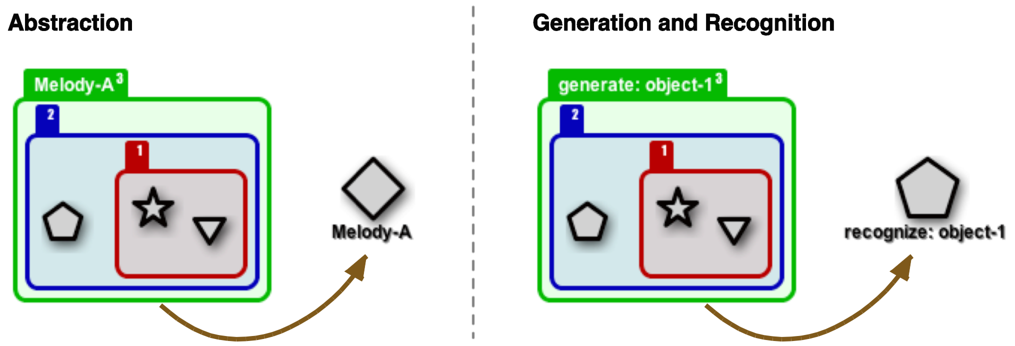


Figure.4 Usage of advanced notations: abstraction, generation, and recognition

Abstraction is a function to encapsulate a structure by defining an arbitrary label and making the structure reusable. Groups whose names start with capital letters, such as “Melody-A” or “Clip-1,” are recognized by the system as Abstraction. A structure grouped as an Abstraction is registered with a label specified by the user. To reuse this Abstraction as a compositional element, the user creates an Atom with the name of the label of the Abstraction. Then the structure is internally reproduced by that Atom. In this way, Abstraction facilitates the creation of an abstract structure class by means of encapsulation.

Generation and *Recognition* are other functions that are used in pairs. In Gestalt Editor, a basic composition is conceptually a set of hidden motifs, but it appears as an existing artifact only when it is *generated*. Then, the user can create a definition to recognize that artifact, and the final result is rendered. Using these notations, the user can simulate how the composition is realized as an artifact, and also how it is recognized as an experience. These features might appear as extra processes compared to other ordinary creation tools. However, we thought it is important to separate clearly *generation* and *recognition* processes from the composition process, because these estimations are usually done on the basis of tacit knowledge, and those parts are exactly what we would like to examine to understand the cognition process.

3.2.4 Designing music and space

The motifs created by Atoms and operations specified by Groups actually correspond to musical and spatial content.

For music creation, Atom creates a motif, which is a sequence of musical notes. A motif and its content have properties such as timing, pitch, duration, timbre, and dynamics. These properties can be logically deformed by an operation. Figure 5 is an example of a musical motif and the operations “reverse” and “inverse.”

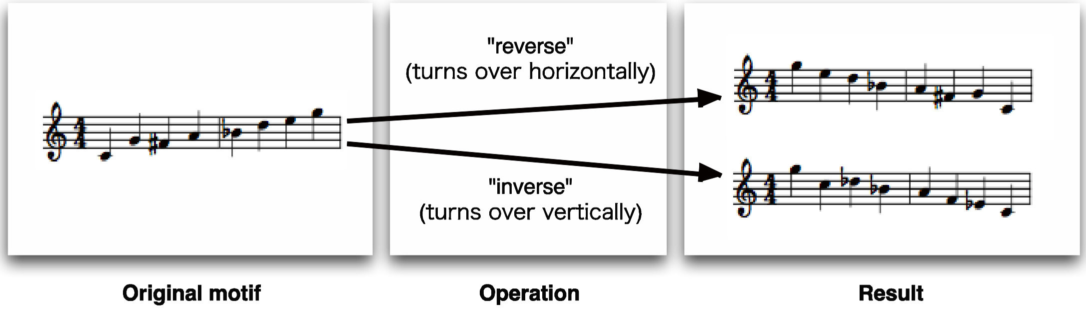


Figure.5 Example of operations for music

For space design, the final output is a slideshow of rendered images (Figure 6). Atom represents a three-dimensional primitive object such as a cube, a sphere, or any other pre-modeled object. As with music, these objects contain properties, but in this case, they are shape, scale, size, location, texture, color, etc. They can be transformed by operations as well. Additionally, using recognition feature, the user can define a location, direction, and scope of the viewpoint.

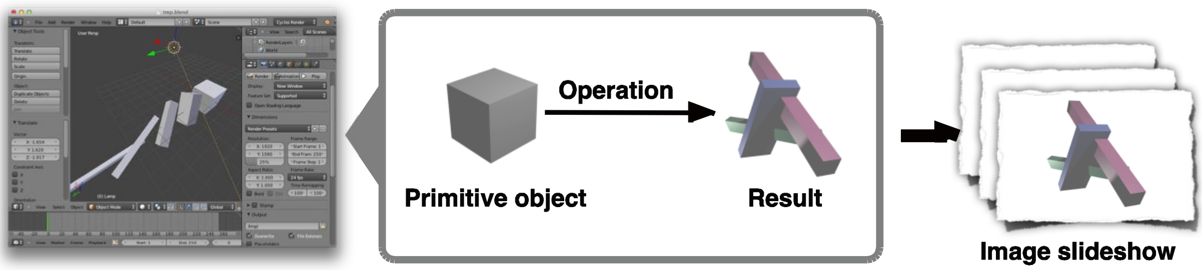


Figure.6 Example of image generation

In order to synchronize music and space, the *x*-directional position of recognition elements is mapped to the timing of playing a motif. Multiple motifs can eventually be combined in a technique similar to counterpoint in music. For space, this is a transitional sequence of an image slideshow.

In the process described above, Gestalt Editor provides an integrated experience of music and space from a user-defined structure.

3.3 Automatic generation of music and space

Figure 7 shows the flow of processes for creating music and space from a user-defined structure and integrating it into a picture slideshow with music. These processes consist of five stages shown in the figure as grey areas [A] to [D]. All these processes are performed automatically. (See Appendix A for the detailed data formats.)

3.3.1 Structure analysis ([A])

First, Gestalt Editor exports the created structure on the canvas to a CSV (comma-separated values) file. The CSV file contains information about all Atoms and Groups with their attributes and relationships.

Next, a program called *Structure analyzer* (denoted by 1 in Figure 7) analyzes the structure data. The structure analyzer serves as an interpreter to trace the structure. This analyzer starts tracing from Atoms representing motif creation, and it follows the route to parent Groups. Based on this tracing result, the order of the actual generation sequence, called a “recipe,” is determined and is written to a text file. Both music and space are generated from the same recipe, which is a result of structure analysis.

3.3.2 Music generation ([B])

Score generator (2) automatically composes musical notes according to the *Generation sequence*. At that time, this generator refers to *Motif data* pre-composed by the user, which is a data set of notes for each motif. Musical operations for motifs are also performed here. The result is exported as a music score in the standard MIDI file (SMF) format. Finally, an external program *Synthesizer* (3) converts the SMF into a sound file.

3.3.3 Space generation ([C])

For space generation, we adopted the open-source 3D graphics software *Blender* (http://www.blender.org/) (5) as the core engine for space generation. Many features in Blender have several advantages beneficial to our purpose. It is capable of both automatic manipulation using scripting and manual editing on graphical user interfaces. It is also possible to edit a virtual space manually after it has been generated by Gestalt Editor.

Since Blender uses Python as the scripting language, we developed a process called *Space preprocessor* (4). This process converts a *Generation sequence* into a Python script. By interpreting this script, Blender (5) constructs a 3D space using external *Model data* and renders the result as an image file.

3.3.4 Integration ([D])

The generated sound and image files are integrated into a picture slideshow. *Web browser* (6) performs this using the recently popular HTML5 and JavaScript library jQuery (http://jquery.com/). Images are synchronized with sound by displaying them in the time sequence specified by the *Recognition* elements in Gestalt Editor. Thereby, the user can experience an integrated output of music and space composed from an identical structure.

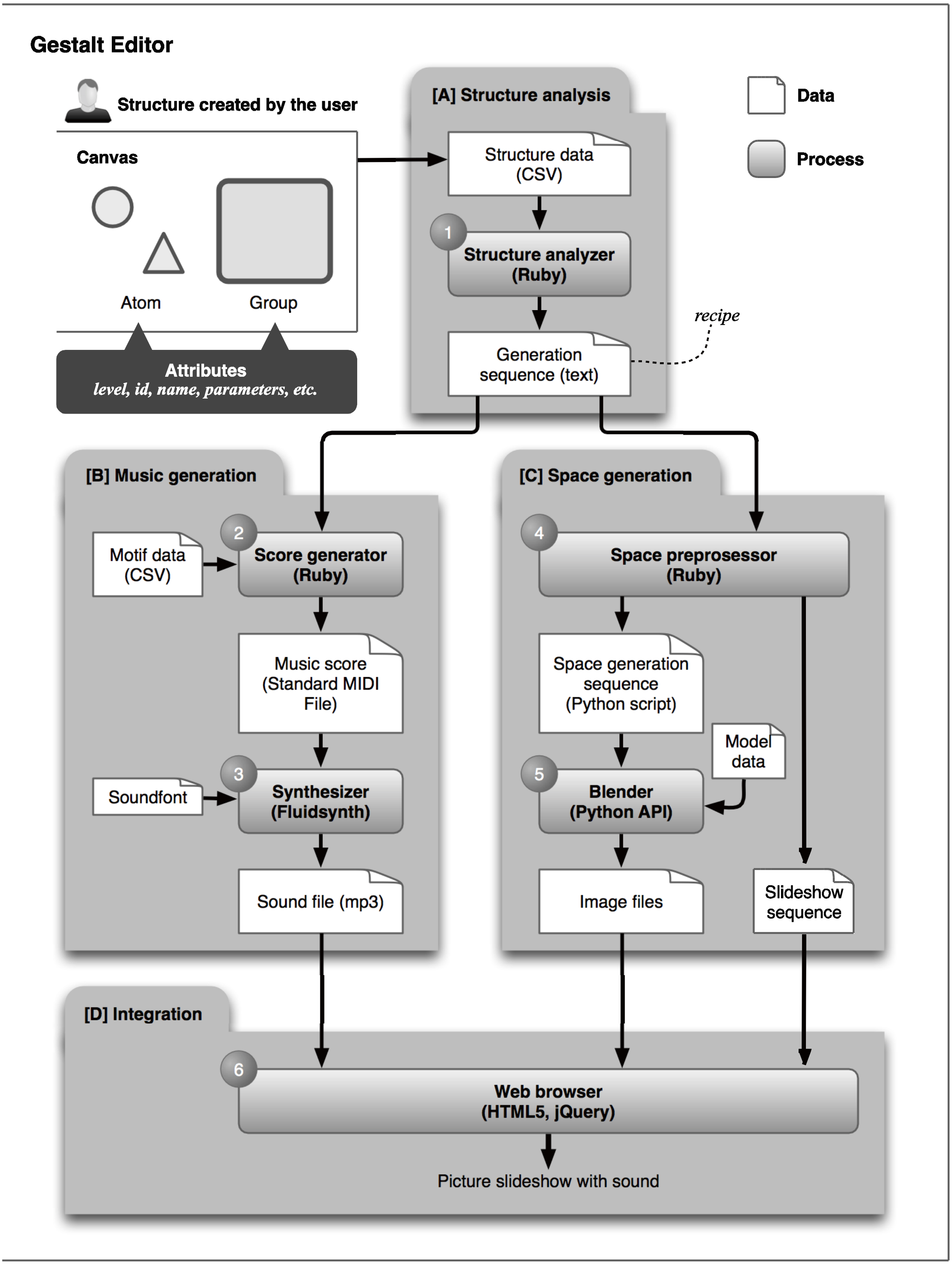


Figure.7 Process flow of automatic music and space generation by Gestalt Editor

4. Discussion

4.1 Usability

Throughout the development of Gestalt Editor, we considered usability as an important factor. Currently, the user can edit a structure on a 2D canvas in a way similar to other drawing software programs. Music and images can then be generated from that structure by a single click. These features contribute to improved usability.

However, there still is room for improving the usability; however, the major problem is the time it takes to generate images. Even though the space composition can be constructed in Blender in only a few seconds, the final image rendering process takes several seconds per image. If the user could check the result much faster using a technique such as real-time rendering, it would make interactivity much smoother.

4.2 Evaluation

We successfully developed an experimental software environment that handles surface structure and deep structure in creative processes and generates music and images. However, in order to confirm the validity of this software environment and its underlying methods, it is necessary to collect sufficient data of composed structures. This process should be done through practical creation by many artists and designers, which has not been done thus far. We are currently planning to release this software environment and to make a community group of users to accumulate more data for analysis. This will lead to further improvements in the software and our method.

5. Conclusion

As a first step of our proposed constructive approach to finding a commonality in the creative process, we developed Gestalt Editor, a software environment for music and space creation that bridges design thinking and creation processes. The result of creation using Gestalt Editor can be applied to research for improving the hypothesized model of deep structure, and that model can also be used in creation. In this way, Gestalt Editor enables a new constructive approach that focuses both on research about the deep structure and the creative process by sharing results between them.

Thus far, we succeeded in creating an integrated experience of music and space from a common structure. However, further evaluation of this method is strongly suggested. We hope that our future efforts will contribute to the elucidation of cognitive processes involved in creative processes and the development of new forms of expression.

6. References

[1] Fujii, H., Furukawa, K., and Kiyomizu, Y. (2006) *Towards a Proposal of Method of Composing and Designing Architectural Integrating Music and Architectural Space,* The 64th IPSJ SIG on Music and Computer, pp. 1-6.

[2] Furukawa, K., Fujii, H., and Kiyomizu, Y. (2006) *Comparative Study on of a Relationship between Music and Space Experience based on Tempo at Circuit Style Garden*, The 64th IPSJ SIG on Music and Computer, pp. 7-12.

[3] Haegeman, L. (1991) *Introduction to Government and Binding Theory,* Blackwell.

[4] Lerdahl, F. and Jackendoff, R.S. (1983) *A Generative Theory of Tonal Music*, MIT Press, Cambridge, Mass.

[5] Ohmura, H., Kimura, R., Fujii, H., Okanoya, K., and Furukawa, K. (2010) *Development of an Editing Tool "Gestalt Editor" to Express Music Multidimensional Structures for Creating Music*, In IPSJ SIG Notes 2010-MUS-87(2), Information Processing Society of Japan, pp. 1-5

[6] Schafer, R.M. (1993) *The Soundscape*, Destiny Books, Rochester, Vt.

[7] Shlomowitz, M. (1999) *Cage's Place in the Reception of Satie*. Available at <http://www.satie-archives.com/web/article8.html>

[8] Stiny, G. and Gips, J. (1972) *Shape Grammars and the Generative Specification of Painting and Sculpture*, In Proceedings of IFIP Congress 1971, North Holland Publishing Co., pp. 125-135

[9] Suwa, M. and Fujii, H. (2009) *Development of a Space-Sound Inter-media as a Tool for Inspiring Meta-cognition in Space,* Annual Conference of Japan Cognitive Science Society 2009, CD-ROM.

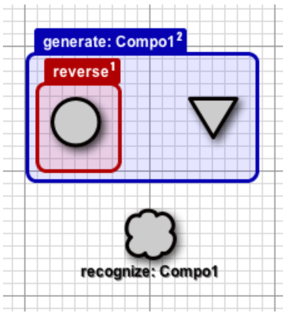
[10] Suwa, M. and Fujii, H. (2010) *Designing a Space-Sound Inter-media as a Tool for Enriching Spatial Experience,* Design Symposium 2010, USB.

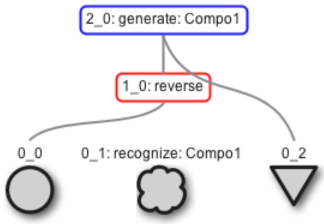
[11] Treib, M. (1966) *Space Calculated in Seconds: The Philips Pavilion, Le Corbusier, Edgard Varèse*, Princeton University Press

Appendix. Example of music and space creation

Below is an actual example of structure composition, data format, and generated music and space (Figure 6).

Original Structure





Structure data (CSV)

TYPE	LEVEL	ID	NAME	PARAMETER	SHAPE/COLOR	X	Y	PARENTS	MY_INDEX	BROTHER	KINSHIP
meta	-1	-1	zoom	40	-1	-1	-1				
meta	-1	-1	grid	true	-1	-1	-1				
meta	-1	-1	snap	false	-1	-1	-1				
atom	0	0			0	136	150	1_0,2_0	0,0	1,2	1,2
atom	0	1	recognize: Compo1		12	138	249				
atom	0	2			2	246	153	2_0	1	2	1
group	1	0	reverse	pattern=+-+	0	136	150	2_0	0	2	1
group	2	0	generate: Compo1		0	136	150				

Generation sequence (text)

```
pc 0 0 0
po 1 0 0
op 2 0 0 -> #0
pc 0 2 0
op 2 0 0 -> #1
om 2 0 0 <- #0 #1
re 0 1 0
```

```
// pc: primitive creation
// po: primitive operation
// op: object - push to array
// om: object - merge from array and register
// re: recognize
```

Generated results

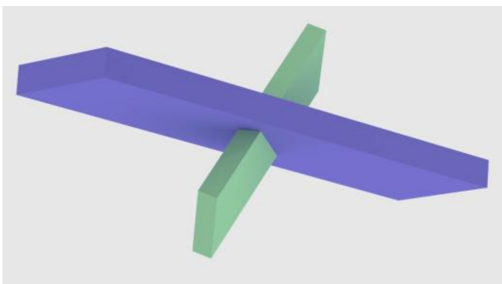


Figure.8 Example of a composed structure and its result by Gestalt Editor