Towards a Coherent Terminology and Model of Instrument Description and Design

Tellef Kvifte
Dept. of musicology
University of Oslo
P.O.Box 1017 Blindern
0315 Oslo, Norway
tellef.kvifte@imv.uio.no

Alexander Refsum Jensenius
Dept. of musicology
University of Oslo
P.O.Box 1017 Blindern
0315 Oslo, Norway
a.r.jensenius@imv.uio.no

ABSTRACT

This paper discusses the need for a framework for describing musical instruments and their design, and discusses some possible elements in such a framework. The framework is meant as an aid in the development of a coherent terminology for describing, comparing and discussing different musical instruments and musical instrument designs. Three different perspectives are presented; that of the listener, the performer, and the constructor, and various levels of descriptions are introduced.

Keywords

Musical instrument design, mapping, gestures, organology.

1. INTRODUCTION

In the literature on musical instruments and musical instrument construction, one central theme is the relation or mapping between gestures[1] used to control an instrument and the resultant sounds. In later years, an increasing number of papers describe aspects of such gesture-sound relationships, many of them basing the discussion on specific examples of newly constructed controllers and/or instruments. In many cases, however, these discussions do not address more general principles, and even if the instruments described are interesting, the discussions do not necessarily add to a broader understanding of musical instrument construction. Part of the problem is a lack of consensus on terminology, and this paper is an attempt to start a discussion of necessary, fruitful and convenient terminology in the study of musical instrument description and construction. Through this, we advocate a field of study which might be called *theoretical organology*.

The construction of new instruments and mappings raises a number of considerations. One set concerns the listener: What kind(s) of sounds do we expect when we see a certain gesture? What kind(s) of gestures do we imagine when we hear certain sounds. Another set concerns the performer: What are intuitive and natural mappings in the instrument? What are exciting mappings that stimulate creativity?

This is not to say that new mappings should always follow such expectations, but a better knowledge of gesture-sound relationships would certainly help in both making mappings that both conform to, *and* possibly also violate, our expectations. In both cases a set of conceptual tools are needed; and to be able to draw on the great number of studies that have been, and will be, published, we need to start discussing to coordinate the terminology being used.

2. THREE PERSPECTIVES

There is of course a close connection between overall perspective and terminology. We have already mentioned two different perspectives, that of the performer and that of the listener. A third is that of the instrument constructor. The following sections will present three models of an overall performer-instrument system corresponding to these three perspectives; and how they can be described.

2.1 The listener

Seen from a listener's perspective, it is important to be able to characterize the general relationship between gestures and the emerging sound (Figure 1).

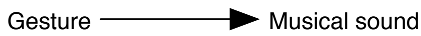

Figure 1: Gesture - Sound

The musical sound may also invoke some kind of listener activity, for example dancing, foot tapping or applause. Figure 2 presents this distinction of Gesture P(erformer) and L(istener).

Figure 2: Musical sound also evoking new gestures

The gestural activity of the audience may in turn influence the musicians and their gestures; and we will have a closed loop of information flow (Figure 3).

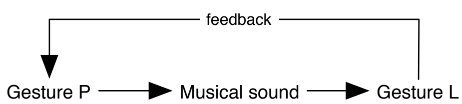

Figure 3: Closed information loop

[1] "Gesture" is here defined as body gestures, i.e. physical movement.

In this paper, where we focus on the description and construction of musical instruments rather than the interplay between musicians and audience, this larger loop is not in the center of attention. We will rather concentrate on how the listener perceives the interplay between the musician, the instrument, and the resulting sound, as seen from a smaller or greater distance. We want to characterize how the actual connection between gestures and musical sound, i.e. the *mapping*, is perceived (Figure 4).

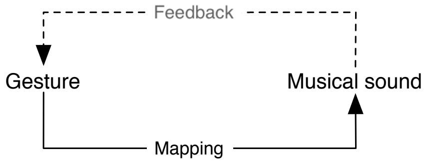

Figure 4: Model of the listeners perspective

Seen from a distance, smaller details of finger movements will probably not be as important as larger bodily movements, and one may not be able to distinguish easily between sound-producing gestures and other movements. Purely expressive and/or optional gestures may nevertheless be experienced as relevant to the sound by the audience, and it is not obvious where to draw the analytical line. As Wanderley et. al observe: "[...] clarinetists' ancillary gestures are not randomly produced or just a visual effect, but rather they are an integral part of the performance process." [10, 98]

To account for this, descriptions of the mapping between gesture and sound at this level might therefore include several levels of resolution.

2.2 The performer

In a model of the performer's perspective, we add a *device* and include the feedback. The performing subject is only implicitely represented in the figure, as the agent performing the Gesture, and receiving the Feedback (Figure 5).

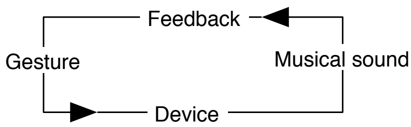

Figure 5: A first approximation of a model of the performer's perspective

Several models building on this basic scheme have been proposed. One rather abstract model is found in [7], and includes just the performer's perceptions and intentions about the playing on the one hand, and what is called the *instrument control* on the other (Figure 6). Implied in the model is that the feedback from the instrument control to the performer is in terms of sound, understood as music. This model is focused primarily on the performer, to the extent that the instrument as such is not present; only the parts of the instrument that are sensitive to control. These are called *control organs* in this model, and we will use this term to denote instrument parts like keyboards, buttons, finger-holes etc., as well as sensors of various kinds that are put to use for controlling sound-producing hard- and software.

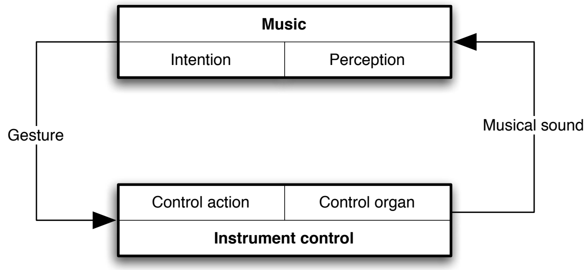

Figure 6: The playing technique perspective after [7]

Choi's model (Figure 7) is a bit more detailed, as the instrument here also includes the sound-producing parts, and the instrument control, here called *interface*, is exemplified to a certain extent. Note that both these models only present auditory feedback; none of them take haptic or tactile feedback into account.

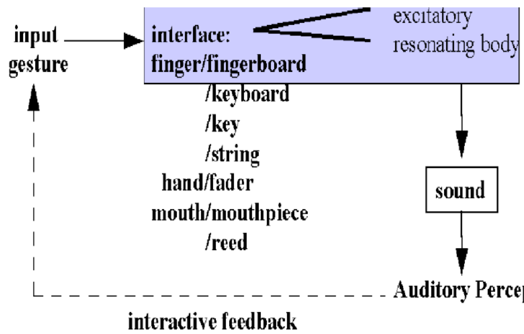

Figure 7: Interactivity of solo performer with musical instrument. After [2]

An even more abstract model is found in [5], as shown in Figure 8. In this model, performer actions and input devices correspond to the control actions and control organs in Figure 6. Here, the sound producing device is included, but the performer and feedback loops are not shown. This model highlights the mappings between performer actions and the sound producing device as a possibly rather complex system of connections.

A more comprehensive model should include the map-

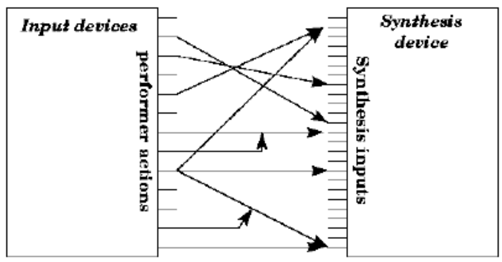

Figure 8: Mapping of performer actions to synthesis parameters. After [5]

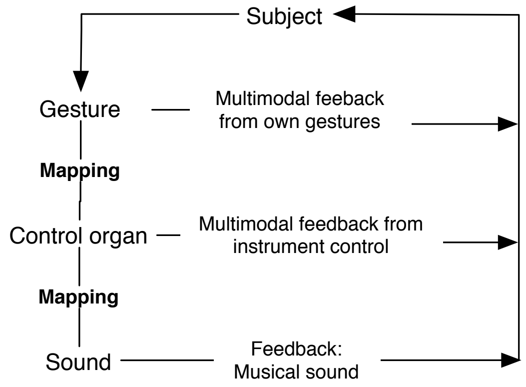

Figure 9: A model of the performer's perspective

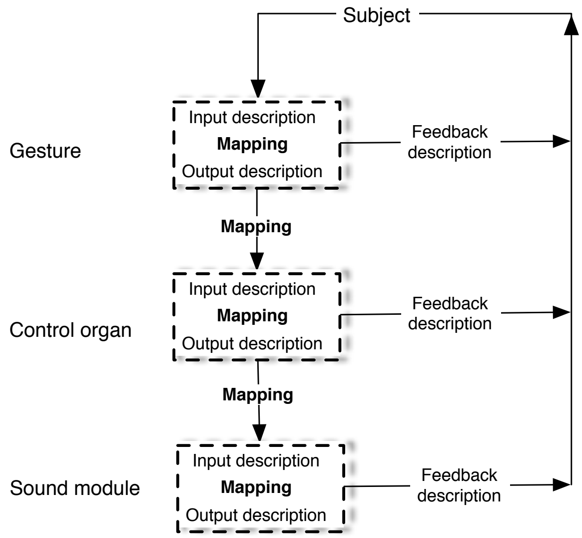

Figure 10: Model of the constructor's perspective

pings "gesture – control organ" and "control organ – sound", but in the musical information flow, the inner workings of the sound module is not of interest to the performer, so that part is reserved for the constructors perspective.

We also find it important to explicitly show the feedback from the different parts of the system. The performer may receive and use feedback from his own gestures, from the contact with the control organs of the instrument, as well as from the actual sound produced (as illustrated in Figure 9). Sometimes the feedback is essential — it is literally impossible to play a theremin without hearing the sound and pitches produced. In other cases feedback may be helpful, but not essential, like the visual feedback from a piano style keyboard, as is illustrated by e.g. professional, blind pianists.

2.3 The Constructor

The constructor needs a far more detailed view of the system than the listener and performer presented in the previous sections. There are several models in the literature, each focusing on different aspects. A relatively comprehensive model might look like the one in Figure 10.

This model suggests that the mappings may be viewed as chained: the output of one mapping is the input to the next. Also, the phenomena to be described are quite diverse; from space-time trajectories (gestures), to interface layout, to sound synthesis descriptors, to descriptions of sound as music.

This model is quite similar to the one found in [1], where the interesting concept of 'related–to–perception' parameters is introduced, (Figure 11), but slightly more explicit in the number of descriptions and mappings. Unlike Arfib's, however, the model in Figure 10 does not say anything about actual parameters to form the descriptions, or about the mappings between them.

2.4 Technical vs. musical construction, or the role of information and energy

In models of the kinds shown here, we will find accounts of both information flow and flow of physical energy, and it is not necessarily obvious whether Figure 10 refers to information or energy, or both.

One reason for this, is that the concept of 'energy' is used in many different ways, both as a concept of physics, but also to describe perceived qualities of music, like when we talk about 'energetic playing', 'forceful sounds' etc.

On closer inspection, it is obvious that the upward flow in the figure refers to information exclusively, and not to physical energy, as the only entities flowing are *feedback descriptions*. This is not to say that physical energy is not involved in the process, but only to point out that what we are interested in, also concerning energy, is the information conveyed.

Striking a piano key requires a performer's energy, and this energy is musically relevant since it determines the force with which the hammer strikes a string, which in turn determines the energy of the sound produced, which in turn determines the energy reaching the ears of the performer and listener. This energy, however, is not the same as perceived loudness, when loudness is seen as a musical parameter. While the physical energy level decreases with the square of the distance from the sound source, the perceived musical loudness may be almost invariant over quite a distance from the sound source. For the performer, the actual energy used is less interesting than the *difference* in energy needed to produce a *difference* in a perceived musical parameter.

In general, as far as Figure 10 is seen as a model of music-making, one should only be concerned with information. The *musical construction of an instrument* is a matter of information processing. However, for the constructor, a piano key must be made so that differences in output levels, corresponding to musically meaningful differences, are physically feasible for the performer. And the chain of energy that carries the information around the circuit must be geared to the equipment used at the different stages.

Such considerations may be kept separate from the musical construction of the system, and we believe the musical construction will become clearer if this separation is made.

3. PARAMETERS

3.1 Parameter types

The parameters involved in the description of a musical instrument may be organized in the following general types:

- Gestural parameters
- Technical parameters

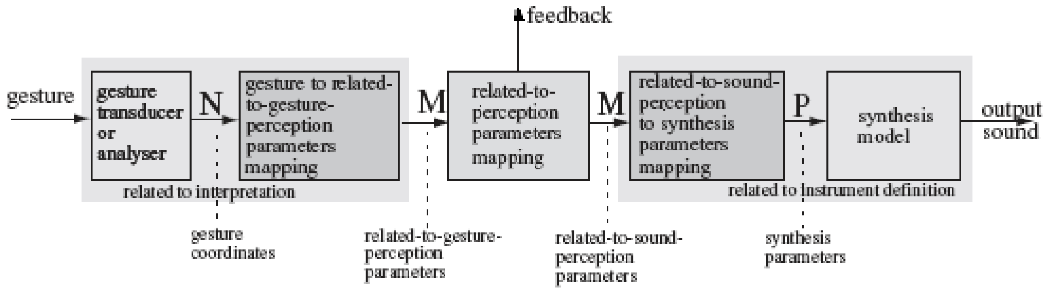

Figure 11: Mapping chain after [1]

- Musical parameters

Typically, the input parameters of the whole system will belong to the first type, while the output will, as argued above, best be described in terms of musical information.

The different types of parameters must be described in different ways, and the challenge is to find ways to connect the descriptions to find common or corresponding properties in the different kinds of description.

3.2 Parameter description

In this section, we will concentrate on properties that are valid across different kinds of parameters and their description in detail. This concerns level of specificity, the distinction discrete/continuous, and the concept of measurement levels. These are all considerations that may be helpful when looking for how properties of parameters at one point are reflected in properties of parameters at another point in the chain.

3.2.1 Levels of specificity

It is obvious that there is a need for descriptions on different levels, or with various degrees of detail. The gestures used by performers might be best described in rather broad terms in the perspective of the listener, while one will need a far greater amount of detail when describing them from the performers perspective and the constructor may need even more detail.

3.2.2 Musical parameters

One of the real challenges, is to define relevant descriptions of musical output. At a very general level, the parameters *pitch, loudness, timbre* and *duration* may be used. As soon as one wants more specific descriptions, however, there is a large number of possible descriptions, and meaningful descriptions are very much depending on musical style, as well as on experience with the instrument in question.

The descriptions needed, develop with experience, and with development of new instruments. From the very start of synthesizer construction, a large number of new parameters, especially for timbre, became available for exploration and incorporation in musical practice through new user interfaces. More recently, a similar development concerning musical manipulation of time in various software packages, widens our musically relevant parameters in the field of rhythm, tempo and time.

In this context, we will touch upon only the most general level.

3.2.3 Levels of measurement

One way to characterize parameters in a general way, is through the concept of *levels of measurement* [4]:

- Nominal level: Values may only be distinguished from each other, and not ordered.
- Ordinal level: Values may also be ordered in a sequence.
- Interval level: Values may be ordered, and there is a way to measure distance between values.
- Ratio level: Values may be ordered; distances measured, and there exists an absolute zero value, so that division and multiplication of values are meaningful operations.

At a general level, these four levels may be associated with the four general musical parameters in the following way:

- Nominal level: Timbre. There is no generally accepted way to characterize and organize timbre, but it is still possible to differentiate between different instruments, and it is possible to construct scales for *aspects* of timbre (e.g. brightness).
- Ordinal level: Loudness. It is obvious how to order loudness levels, but not necessarily how to describe precise *intervals* of loudness.
- Interval level: Pitch. Musical pitches are ordered into classes with perceptually meaningful comparable distances, like half tone, whole tone etc. But there is no meaningful zero point.
- Ratio level: Durations may be ordered and measured, and there is an obvious zero (no duration), and ratios are meaningful — a half note is half the duration of a whole note.

The implication for the control of musical instruments is that an input parameter has to be on at least as high a level of measurement as the output parameter it is meant to control. But there is not always a need for controlling an output variable at the maximum level, as will be explained in more detail below.

3.2.4 Discrete vs. continuous parameters

The perception of the four general musical parameters may be described as having two different dimensions, *discrete* and *continuous*. The discrete dimension is tied to categorization or sets of concepts; the continuous to gradual variations. Pitch may be perceived both as belonging to discrete pitch-classes (c, d etc), *and* as a continuous entity as in vibrato and glissandi. Similar considerations may be done for the three other general parameters (See e.g. [3] and [7]).

There are several implications of this for the connection between input (gestural) parameters and output (musical) parameters (Figure 12). First of all, a continuous output parameter needs a continuous input control to be controlled in detail. But it is also possible to use a discrete input to *trigger* a preprogrammed continuous variation of output like a vibrato on an synthesizer. On the other hand, a discrete output may be controlled by a discrete input, like pitch classes controlled by a keyboard, but there are also numerous examples of a continuous controller controlling a discrete output parameter, like a trombone slide controlling discrete pitch classes (as well as continuous pitch variations).

Input \ Output	Discrete	Continuous
Discrete	Piano keyboard: discrete pitch	LFO triggers in synths. Rare in acoustic instr.
Continuous	Trombone slide: discrete pitch	Trombone slide: continuous pitch Striking force: loudness

Figure 12: Combinations of discrete and continuous input and output

The various combinations of discrete and continuous variations for input and output for different musical parameters give different demands and possibilities for the performer. This is generally acknowledged by performance teachers and students, even though we have little general knowledge on the effects of different combinations.

3.3 The description of mappings

3.3.1 The mapping chain

In the three perspectives discussed above, the mapping chain was described in increasing detail; with only one mapping from gesture to sound in the listeners perspective, to a chain of five different mappings in the constructors perspective.

All the mappings in the last chain will at some point have to be described during the construction of an instrument, but they will in a sense be subordinate to the overall gesture–sound mapping, that might be called the *defining mapping* for the instrument. It is this overall gesture–sound mapping which defines the identity of an instrument, and this is what we will focus on in the following sections.

3.3.2 Basic mapping strategies

There are three basic approaches to the actual mapping. The traditional way is to describe and construct a fixed or static mapping, where the relations between input and output parameters stay the same. Traditional acoustic instruments are usually well suited to such descriptions.

Another possibility is a variable mapping, where mappings may be changed by the performer. Most commercial synthesizers are good examples of this approach.

Finally, mappings may be the outcome of a dynamic (learning) process where the performer can choose the gestures, or the mapping is modified by the actual behaviour of the performer (e.g. [1] and [9]).

All three cases need a common way of describing mappings.

3.3.3 A general mapping description problem

Several authors mention that mappings may take different forms with regard to how many input parameters are controlling how many output parameters. In [7], the term ‘coupling’ is used, differentiating between ‘control couplings’ and ‘sound variable couplings’, corresponding to the many-to-one and one-to-many examples in Figure 13 respectively (see also [1] and [5]).

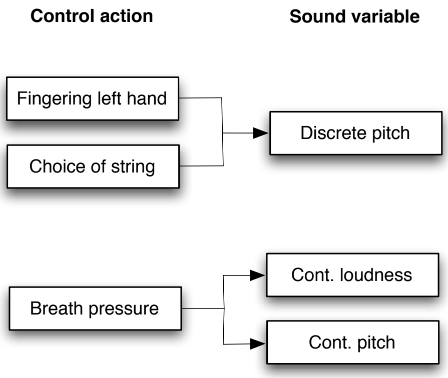

Figure 13: Examples of many-to-one and one-to-many mappings of input to output.

While there are a number of high-level characterizations available, there is a need for a more detailed typology of mappings. Traditional acoustical instruments may be a good starting point, because they represent a quite diverse set of possibilities that are well known through long practice.

One way is to make general overviews of mappings for a number of instruments, where control actions are mapped to the generalized musical parameters as described above. In [7], this is done for 23 instruments, and a few general points are made (one example in Figure 14).

With such overviews, it might be possible to create ‘gestural ensembles’ based on the gestures used to control an instrument. ‘Bowing’, ‘string-stopping’, ‘finger-hole-fingering’, ‘piano-type-keyboarding’, ‘concertina-buttoning’ etc. might be examples of such ensembles. Such entities will represent a specification of the control of single instruments, and at the same time be useful over a range of instruments; possibly useful also as a basis for a more general typology of gesture–sound mappings in musical instruments.

Also relevant to such a typology, is the more detailed model found in [8], where an analysis of how a single musical tone — a musical *event* — is controlled, is used to develop terminology for the description of what is called ‘musical control space’. Here, a musical event is broken

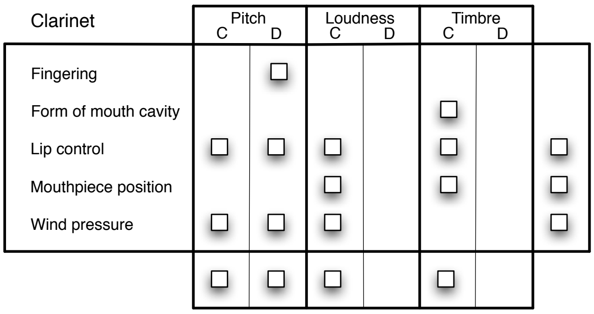

Figure 14: Mapping chart for clarinet. C and D means Continuous and Discrete (see text for explanation)

down into five phases, called *Selective preconditioning, Beginning, Middle, End* and *Terminus*, and analytical categories introduced for each phase.

All these approaches, however, raise a general problem. While input parameters may be relatively well-defined gestures, the same is not the case with output parameters. As discussed in section (3.2.2), meaningful descriptions of musical output in more detail depends heavily on musical style and experience with the musical instrument being discussed.

This means that development of a more general mapping theory should be closely connected to development in *music theory*. Without some kind of consensus on relevant musical entities, what they are, and how they are related, no coherent theory of mappings is possible.

4. CONCLUSIONS

As is obvious from the brief literary overview in this paper, there is no general agreement on terms for describing musical instrument design and mapping. This holds on almost every level of description, and almost every aspect of the models presented here. We believe that a common set of descriptors could be valuable for both analysts and constructors in the field, and may also help to unify various models being proposed, as well as clarifying real differences of opinion.

This paper has not addressed questions of how gestures can be formalized. This is discussed further in in [6].

5. REFERENCES

[1] D. Arfib, J.-M. Couturier, L. Kessours, and V. Verfaille. Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7(2):135–152, 2002.

[2] I. Choi. Gestural primitives and the context for computational processing in an interactive performance system. In M. Wanderley and M. Battier, editors, *Trends in Gestural Control of Music*, pages 139–172. Ircam - Centre Pompidou, 2000. 352.

[3] E. F. Clarke. Categorical rhythm perception and event perception. In *Proceedings of the International Music Perception and Cognition Conference*. Keele University, Department of Psychology, 2000.

[4] O. Hellevik. *Forskningsmetode i sosiologi og statsvitenskap*. Universitetsforlaget, Oslo, 1971.

[5] A. Hunt, M. M. Wanderley, and R. Kirk. Towards a model for instrumental mapping in expert musical interaction. In *Proceedings of the 2000 International Computer Music Conference*, pages 209–212. International Computer Music Association,, San Fransisco, 2000.

[6] A. R. Jensenius, T. Kvifte, and R. I. Godøy. Towards a gesture description interchange format. *NIME06, Paris*, Fortchoming.

[7] T. Kvifte. *Instruments and the Electronic Age.* Solum Forlag, Oslo, 1989.

[8] D. J. Levitin, S. McAdams, and R. L. Adams. Control paramteres for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(2):171–189, 2002. 344.

[9] D. J. Merrill. *FlexiGesture: A sensor-rich real-time adaptive gesture and affordance learning platform for electronic music control.* Master of science in media arts and sciences, MIT, 2004.

[10] M. Wanderley, B. W. Vines, N. Middleton, C. McKay, and W. Hatch. The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113, 2005.