On the Choice of Transducer Technologies for Specific Musical Functions

Marcelo M. Wanderley, Jean-Philippe Viollet, Fabrice Isart, and Xavier Rodet

Ircam – Centre Pompidou
Analysis-Synthesis Team
Paris – France

mwanderley@acm.org, rod@ircam.fr

ABSTRACT

This paper presents a study of gestural control applied to sound synthesis. It discusses an implementation of real-time gestural control of the CHANT software (Max/MSP). The input device consists of a Wacom graphic tablet instrumented with additional sensors. A study has been carried out on the selection of transducer technologies to match the available synthesis parameters, based on existing literature. In order to prove the validity of previous studies, different user tests have been devised and implemented. User reactions were annotated and further evaluated. An exploratory analysis of the tests' results is presented which tends to validate previous works.

1. INTRODUCTION

The choice of input device to perform the control of sound synthesis is an important task for instrument designer. One can either propose new devices from scratch or use existing designs, depending on the specific constraints in each context[1]. Real-time control of sound synthesis by means of alternative input devices presents the advantage of providing the user with new possibilities regarding the gestural vocabulary to be used.

The design of input devices for musical use has been the theme of previous studies, such as [Car92] [Car94] [Chu99] and [Bon00]. Similarly, the evaluation of existing input devices for musical tasks was carried out in [VEC94][Ver94], and [VE96].

One alternative input device (or controller) that seems to be gaining increasing acceptance in the musical domain is the graphic tablet. One early example of its use has been shown in [RM88]. Specifically, the Wacom graphic tablet provides the user with many simultaneous degrees of control [WWF97]. When instrumented by the addition of extra sensors [SDWR99], its capabilities may be augmented in order to fit specific contexts.

The question that nevertheless remains is how to select the best transducer technology to match the input parameters of a given synthesis algorithm.

2. CONTEXT OF THE STUDY

The goal of this paper is to study real-time control of CHANT (Max/MSP) using a Wacom graphic tablet as the main input device.

In this section we will review notions related to performer actions, the application concerned and finally the input device.

2.1 Performer Actions

This research is based on an analysis of performer actions and their function in a musical context. We shall consider '*action*'[2] as referring to any body movement, be it free or applied to a tool, that allows a human to interact with his environment.

Hand gestures and hand gesture functions

Cadoz [Cad94] associates to the hand 3 different functions:

- *The Ergotic Function*: the action of modifying and transforming the environment.
- *The Epistemic Function*: the action of getting knowledge from the environment.
- *The Semiotic Function*: the action of conveying information to the environment.

[1] Part of this question can be answered by the selection of appropriate mapping strategies relating the input device parameters to the synthesis ones [HWK00].

[2] The word *gesture* is sometimes used in the literature instead of *action*.

Instrumental gesture definition

An instrumental gesture is an action applied to a material object, where physical interaction takes place. Within this interaction, specific physical phenomena are produced, whose forms and dynamical evolution can be controlled by the one applying the gesture. Moreover, these phenomena can convey communicational messages (information) or are the basis for the production of a material action [CW00].

Instrumental gesture typology

Cadoz [CW00] proposes an instrumental gesture typology according to the gesture's function as follows:

- *Exciter gesture*: that which conveys the energy that will be found in the sonic result. It can be *continuous*, *instantaneous* or *sustained*.
- *Modification gesture*: that which modifies the properties of the instrument but whose energy does not participate directly in the sonic result. These can be divided into two groups :
  - *Parametric modification gesture* - continually changes a parameter.
  - *Structural modification gesture* - modifies the structure of the object (instrument).
- *Selection gesture*: that which performs a choice among different but equivalent structures to be used during a performance.

2.2 Application concerned

The synthesis software used is CHANT [RPB84], originally developed at *Ircam* in the 1970s. It is based on *formant wave forms (FOF)*.

In this synthesis model, each formant is synthesized separately for each period of the pseudo-source and then summed with the other formants. Each formant corresponds to one resonant filter and its main parameters are *center frequency, bandwidth, and amplitude***.** The *fundamental frequency* is given by a sequence of impulses, which is sent into these resonant filters.

All these parameters have been implemented in Max/MSP [Zic98] by F. Iovino and R. Dudas. This implementation also takes into account the *general amplitude* and *vibrato*.

2.3 Input Device used

In our study, the main input interface consists of a Wacom ArtZ II digitizing tablet, with both a stylus (pen) and a puck (mouse-like device).

The choice of this device is based on the available control features of the tablet. In addition to the 2 dimensional position of the stylus and the puck, the tablet also outputs the tilt (angle) in both X and Y directions, and pressure values of the stylus. The puck has four buttons and the stylus itself has two buttons.

We instrumented the tablet with extra sensors that measure only force (a round FSR[3]) or position and force simultaneously (a linear FSR). We have placed them on the left side[4] of the tablet in order to allow extra input variables (such as finger pressure)[5].

The question is how best to relate these control possibilities to the synthesis parameters.

3. MATCHING TRANSDUCER, FEEDBACK AND MUSICAL FUNCTION

Vertegaal et al. [VUK96] [UV00] proposed a simple three-tier classification of musical functions and relate it to the type of transducer[6] and kind of feedback. The proposed musical functions were the following:

- *Absolute dynamical functions*: absolute selection of a pitch or amplitude value.
- *Relative dynamical functions*: modulation of a given pitch.
- *Static functions*: the selection of pitch range, duration or transposition, etc.

The available transducer types analyzed were:

- *Position* (linear or rotary);
- *Velocity* (linear or rotary);
- *Isometric force* (linear or rotary);
- *Isotonic force* (linear or rotary);

In order to give an example of the use of this classification, one can see from [VUK96][7] that an isometric force transducer[8] shall perform well for a relative dynamical function and give excellent tactile and kinesthetic feedback, but no visual feedback.

4. GESTURAL CONTROL OF CHANT

In order to establish a real-time control of CHANT, we related the tablet variables with the classification proposed in [VUK96]. This results in the following correspondence:

Input Devices	***Transducer Type***
Stylus X, Y	position
X-tilt or Y-tilt	rotary position
buttons	position
pressure	isometric force
Puck X, Y	position
buttons	position
Extra Sensor pressure	isometric force
Extra Sensor position	position

Table 1: Transducer types available.

At first, we established the control of the CHANT synthesis by using pre-selection values for each of the 5 FOF parameters in order to simplify the cognitive load of simultaneously controlling all synthesis parameters from scratch. In fact, it is more intuitive to find a pertinent control with less synthesis parameters, since the parameters of the five FOFs used in this program are not directly linked to the others[9].

In order to find the best transducer for each of these three parameters, we related them to the type of musical function and instrumental gesture. We thus obtain the following table:

	Musical Function	***Gesture Type***
Fundamental frequency	absolute dynamical	selection / parametric modification
Vibrato	relative dynamical	parametric modification
General Amplitude	absolute dynamical	excitation

Table 2: Musical Function and instrumental gesture type associated with the synthesis parameters.

Using these remarks, we selected from [VUK96] the transducer type that best fits the musical function. This gives the following:

Musical Function	***Best Transducer type***	***Second transducer type***	***Third transducer type***
Absolute dynamical	position	isometric force	isometric force *(rotary)*
Relative dynamical	isometric force	isotonic force (rotary or not) or velocity	position *(rotary or not)*
Static	position *(rotary or not)*	isotonic force *(rotary)*	-

Table 3: Musical function and resulting transducer type.

This example shows that a pertinent *one-to-one* mapping can be, for instance:

X position of the stylus	Fundamental frequency
Pressure sensor	Vibrato
Pressure of the stylus	General amplitude

Table 4: Match between transducer and musical function.

5. EXPERIMENTS

The relationship obtained above is a theoretical result. Therefore, we verified it by a practical experiment.

5.1 Initial Considerations

Our initial approach consisted in adapting standard HCI (Human-Computer Interaction) user tests, such as pointing to a target or moving the cursor between two or more targets and evaluating the obtained users' results. One difference of our approach versus standard HCI tests consists in the evaluation procedure, since a quantitative evaluation, such as performance time or accuracy would not necessarily be meaningful in a musical context. We chose a qualitative subject evaluation of each transducer after each test, and also a general subject evaluation at the end of the test.

Another difference from HCI tests refers to the selection of the task, in order to maximally isolate the variable of interest from other parameters. Since musical performance is typically a multivariable task, it is important to consider how best to isolate the variable of interest without losing a minimal musical context. In our case, we are interested in the modulation effect, and so there is no need to evaluate the user accuracy in selecting a specified pitch. On the other hand, we considered that with the simple application of a modulation to one single (unique) note, the test would lack musical context. Therefore, we defined a test where movement is required to select a note, in order to reproduce a basic musical task of selecting one pitch, before applying a modulation to it.

5.2 Test definition

We performed initial tests where the layout of notes from a piano keyboard was mapped onto the tablet's surface, from where the subject had to select the correct pitch from a melody[10] and subsequently apply the modulation. Although this pre-test matched more closely a traditional musical context, it was decided that the cognitive load of asking the subjects to learn how to play the melody by choosing the appropriate keys with the stylus could interfere in our evaluation.

---

[3] FSR: force sensing resistor.

[4] But can easily be moved to the right for a left-handed user, for example.

[5] A figure displaying all the control possibilities with this set-up is available in [SDWR99].

[6] In this paper, the words *transducer* and *sensor* are used interchangeably, although the former more accurately denotes the general principle and the latter the device.

[7] A copy of this paper can be obtained from: *http://www.cs.queensu.ca/~roel/publications/ICMC96/paper.html*

[8] The difference between « *isometric* » and « *isotonic* » force transducer is that with « *isotonic force transducers* », motion is needed to operate the sensor. With « *isometric force transducers* », no motion is needed [VUK96].

[9] The five FOF parameters account for timbral properties that were not addressed in the test.

[10] The song « *Frère Jacques* ».

In order to simplify the task presented to the subjects, we defined four small squares on the tablet, each generating a specific fundamental frequency. The subjects were asked to play these four notes twice, and the clockwise movement through these four rectangles produced the melody.

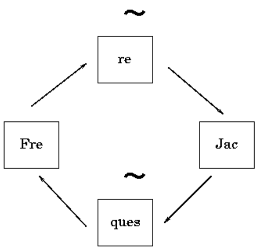

Figure 1: Marking on the tablet's surface - initial experiments.

Results from these initial tests were non-conclusive. For a detailed review refer to [VIWR99] or [Vio99].

5.3 Experiment

In the present experiment, we used three different transducers: *a pressure sensor (round FSR), a position sensor (linear FSR)* and *the X-Tilt of the stylus (Angle X).*

The relative dynamic function used is based on the production of the pitch modulation directly with a finger, hand or wrist movement. For example, when pressing the (pressure) sensor, a number corresponding to the sensor's output value is added to the fundamental frequency, and by changing the sensor's value, the fundamental frequency is modulated. The same principle is used to create this modulation with the position sensor or the X-Tilt of the stylus (one really has to move the stylus back and forth in order to hear a modulation of pitch).

In order to concentrate on the question about the type of transducer, the trajectory was defined as a simple two-way movement between two squares representing only two notes. In order to see whether the use of the X-Tilt depended on the direction of the stylus movement, we proposed two different directions: *a horizontal and a diagonal one.*

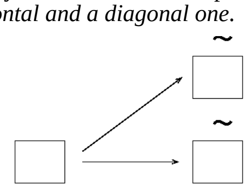

Figure 2: Marking on the tablet's surface - final experiment.

Twenty subjects performed this experiment. Sixteen of them had previous experience in music, but none had significant experience using the tablet. They were asked to estimate the quality of the modulation performed with each transducer.

Each subject listened to a presentation of the task, then to two target sound samples, which were played to give an idea of the tempo, and of the sound effect (a horizontal one and a diagonal one). The subject was allowed a training period of one minute per controller, after which he listened once more to the two target samples and then started the test. Each subject performed the task three times.

5.4 Results

The results shown in figures 3 and 4 correspond to user evaluation for the third performance of the test[11], the first two being considered as learning period and discarded.

As predicted by the theoretical study presented in section 4, users clearly preferred the pressure sensor (round FSR) to perform the modulation.

[11] To see which controller is the best ranked after the test, we use the following weighting to the different criteria: excellent: 7, very good: 6, good: 5, reasonable: 4, bad: 3, very bad: 2, null: 1.

It is interesting to notice that the rotary position sensor (angle X) ranked clearly below the linear position one (linear FSR), a point that was not predicted by the theory. It can be partly explained by the fact that the test is performed using both hands for both the pressure and the linear position sensors, whilst for the Tilt Angle the absolute selection and the modulation were performed with the same hand[12]. This point needs to be further evaluated in order to check whether it is only a result of the test setup or effectively these transducers present different matches, what would contradict the information proposed in [VUK96].

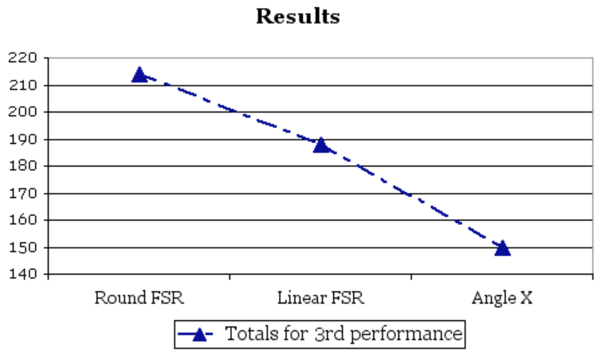

Figure 3: Results summing both directions.

Analyzing figure 4, one can see that the theoretical prediction concerning the pressure sensor was verified for both directions independently. The diagonal direction slightly outperformed the horizontal one for the pressure sensor and both were equally ranked for the position sensor. On the other hand, there is an improvement in the horizontal direction when compared to the diagonal one for the X-tilt sensor. This could be explained by the better integration of the two movements (reaching the square and making the modulation) in the horizontal direction [JSMM94].

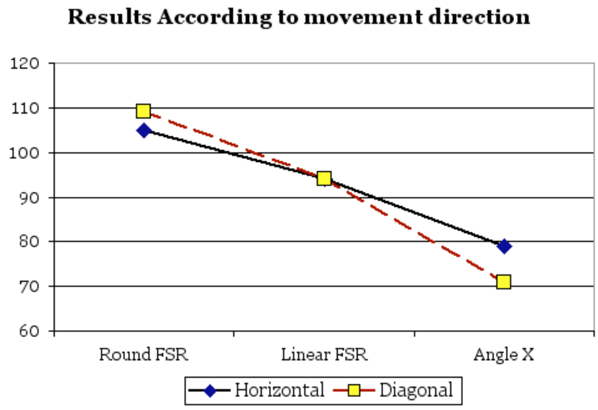

Figure 4: Results for each direction.

6. APPLICATION EXAMPLE: A GENERAL CONTROL OF CHANT

As an example, we propose a general control of Chant synthesis including the 5 FOF parameters. We have chosen, for the fundamental frequency, the vibrato and the general amplitude, respectively, a *position* sensor, a *pressure* sensor and the *stylus pressure*.

But in order to have the whole tablet surface available to control the FOF parameters, we propose to control the fundamental frequency with the position sensor instead of the X tablet dimension. The modulation can then be controlled by the pressure applied to this transducer, since it is both a position *and* pressure transducer.

[12] Nevertheless, in previous tests with a similar setup (cf. section 5.2), we could not find significant differences among the transducers.

It is common to represent the two first formant values of vowels using a triangle diagram [Str65]. The first level of this mapping relates the transducers' outputs to the X and Y coordinates of the triangle [WSR98], which define the triangle used. The second level maps the triangle to the synthesis parameters, *cf.* fig. 5.

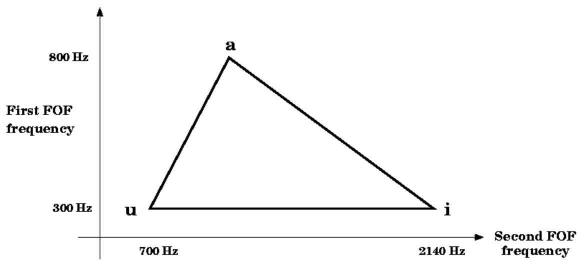

Figure 5: Representation of the first two FOF frequencies on the tablet's surface.

The three other formants are less significant in the differentiation of the vowels. In fact, these values are quite constant for French vowels, and we keep them constant in our mapping (values chosen using [BR89]):

- *Third formant freq.: 2850 Hz*
- *Fourth formant freq.: 3800 Hz*
- *Fifth formant freq.: 4950 Hz*

In the same way, the amplitude and bandwidth values are not significantly different from one vowel to another, and thus we also keep them constant. It would be possible, however, to modify the bandwidth values using the X and Y tilt of the stylus. This may also be interesting because of the integration of different gestures, i.e., to control the first bandwidth parameter with the Y-Tilt and the second one with the X-Tilt, since the first FOF frequency is controlled by the Y tablet dimension and the second one by the X dimension.

7. CONCLUSIONS

In this article we have studied the choice of different transducers for the implementation of specific musical functions. The validity of previous works relating a specific musical function to a type of transducer was discussed. Theoretical results were confirmed by an exploratory data analysis of user tests concerning the choice of transducer technology to perform a specific musical function.

The definition of the test methodologies had to be specially adapted for the musical context. This is an interesting question, since it is not obvious how far can one adapt existing methodologies from other fields. Our approach was inspired by standard HCI methodologies, but ensuring a minimal "musical" context. The evaluation of the proposed methodology had also to be considered in the light of musical performance, in opposition to accuracy and/or time measurements in HCI methodologies.

Finally, we expect this study as a contribution to the discussion of new methodologies for the evaluation and/or further improvement of existing input devices in the context of semi-expert and expert manipulation tasks.

8. ACKNOWLEDGMENTS

We would like to thank all the subjects who participated in our experiments, P. Pierrot and A. Terrier for the extra sensors, and the Ircam Psycho-Acoustics Team for advice concerning the elaboration of the tests.

Thanks to Roel Vertegaal, Suzanne Winsberg and Andy Hunt for various comments and suggestions. Thanks also to Marie-Hélène Serra for welcoming the third author during his period at Ircam.

This research is partly supported by a grant from the France-Berkeley Fund and from CNPq/Brazil.

9. REFERENCES

[BR 89] Bennett, G and Rodet, X. 1989. "Synthesis of Singing Voice." – In Max V. Mathews and John R. Pierce, eds. *Current Directions in Computer Music Research*, MIT Press.

[Bon00] Bongers, B. 2000. "Physical Interfaces in the Electronic Arts. Interaction Theory and Interfacing Techniques for Real-time Performance." In M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*. Ircam-Centre Pompidou.

[Cad94] Cadoz, C. 1994. "Le geste canal de communication homme-machine. La communication instrumentale." *Sciences Informatiques* - Numéro Spécial : Interface Homme Machine 13 (1): 31-61.

[CW00] Cadoz, C. and Wanderley, M. 2000. "Gesture - Music." In M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*. Ircam – Centre Pompidou.

[Car92] Cariou, B. 1992. "Design of an Alternate Controller from an Industrial Design Perspective." In *Proc. ICMC'92*, pp. 366-367.

[Car94] Cariou, B. 1994. "The aXiO MIDI Controller." In *Proc. ICMC'94*, pp. 163-166.

[Chu99] Chu, L. 1999. "MusiCloth: A Design Methodology for the Development of a Performance Interface." In *Proc. ICMC'99*, pp. 449-452.

[HWK00] Hunt, A., Wanderley, M., and Kirk, R. 2000. "Towards a Model for Instrumental Mapping in Expert Musical Interaction." In *Proc. ICMC'2000* (this volume).

[JSMM94] Jacob, R., Sibert, L., McFarlane, D., and. Mullen Jr., M. 1994. "Integrality and Separability of Input Devices." *ACM Transactions on Human-Computer Interaction* 1(1): 3 -26.

[RM88] Raczinski, J.-M., and Marino, G. 1988. "A Real-Time Synthesis Unit." In *Proc. ICMC'88*, pp. 90-100.

[RPB84] Rodet, X., Potard, Y., and Barrière, J.-B. 1984. "The CHANT Project: From the Synthesis of Singing Voice to Synthesis in General." *Computer Music J.* 8(3) : 15-31.

[SDWR99] Serafin, S., Dudas, R., Wanderley, M., and Rodet, X. 1999. "Gestural Control of a Real-Time Physical Model of a Bowed String Instrument. " In *Proc. ICMC'99*, pp. 375-378.

[Str65] Straka, G. 1965 *Album Phonétique*. Les Presses de l'Université de Laval.

[UV00] Ungvary, T. and Vertegaal, R. 2000. "Cognition and Physicality in Musical Cyberinstruments." In M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*. Ircam – Centre Pompidou.

[VEC94] Vertegaal, R. Eaglestone, B. and Clarke, M. 1994. "An Evaluation of Input Devices for Use in the ISEE Human-Synthesizer Interface." In *Proc. ICMC'94*, pp. 159-162.

[Ver94] Vertegaal, R. 1994. "An Evaluation of Input Devices for Timbre Space Navigation." *Master's thesis*. Department of Computing - University of Bradford.

[VE96] Vertegaal, R., and Eaglestone, B. 1996. "Comparison of Input Devices in an ISEE Direct Timbre Manipulation Task." *Interacting with Computers* 8(1):13-30.

[VUK96] Vertegaal, R., Ungvary, T., and Kieslinger, M. 1996. "Towards a Musician's Cockpit: Transducers, Feedback and Musical Function." In *Proc. ICMC'96*, pp. 308-311.

[Vio99] Viollet, J.-P. 1999. "Etude de Dispositifs de Contrôle Gestuel de la Synthèse Sonore." *Technical Report*. École Nationale Supérieure des Télécommunications de Paris.

[VIWR99] Viollet, J.-P., Isart, F., Wanderley, M., and Rodet, X. 1999. "On the Choice of Transducer Technologies for Specific Musical Functions. A Study on Gestural Control Applied to Sound Synthesis." *Int. Report*. Ircam – C. Pompidou.

[WSR98] Wanderley, M., Schnell, N., and Rovan, J. 1998. "ESCHER: Modeling and Performing "Composed Instruments" in Real-time" In *Proc. IEEE SMC'98*, pp. 1080-1084.

[WWF97] Wright, M., Wessel, D., and Freed, A. 1997. "New Musical Control Structures from Standard Gestural Controllers." In *Proc. ICMC'97*, pp. 387-390.

[Zic98] Zicarelli, D. 1998. "An Extensible Real-time Signal Processing Environment for MAX." In *Proc. ICMC'98*, pp. 463-466.