HandSketch Bi-Manual Controller

Investigation on Expressive Control Issues of an Augmented Tablet

Nicolas D'Alessandro
Laboratoire de TCTS - FRIA/FNRS
Faculté Polytechnique de Mons
31, Boulevard Dolez - Mons, Belgium
nicolas.dalessandro@fpms.ac.be

Thierry Dutoit
Laboratoire de TCTS
Faculté Polytechnique de Mons
31, Boulevard Dolez - Mons, Belgium
thierry.dutoit@fpms.ac.be

ABSTRACT

In this paper, we present a new bi-manual gestural controller, called HandSketch, composed of purchasable devices : pen tablet and pressure-sensing surfaces. It aims at achieving real-time manipulation of several continuous and articulated aspects of pitched sounds synthesis, with a focus on expressive voice. Both prefered and non-prefered hand issues are discussed. Concrete playing diagrams and mapping strategies are described. These results are integrated and a compact controller is proposed.

Keywords

Pen tablet, FSR, bi-manual gestural control.

1. INTRODUCTION

As explained in [11], the development of gestural controllers for sound synthesis is today a bit less a pioneer's business. Indeed, plenty of sound synthesis modules are now available in usual and inexpensive computers. Integrated toolkits, with various sensors, user friendly analog-to-digital converters and configuration softwares, can be purchased at affordable prices. Moreover, several environments for mappings and synthesis implementation, such as Max/MSP, and powerful transmission protocols (e.g. OSC) are well supported and integrated on usual plateforms.

Thus, in several contexts, the matter is less having technical answers, than formulating good questions. We can observe that concretization of an original musical controller is often part of the compositionnal process. This vision can easily be extended to scientific research. Particullary, in signal processing, improvements can be realized thanks to real-time manipulations and observations of results.

In this project, we aim at developing a gestural controller dedicated to real-time voice processing in some stationnary and articulative aspects. Section 2 explains technical choices related to pen-based continuous pitch and intensity control. Section 3 proposes an embedded controller for articulated finger-based gestures. Both approaches are joined together in order to propose the "HandSketch" controller. A typical playing situation is illustrated in Figure 1.

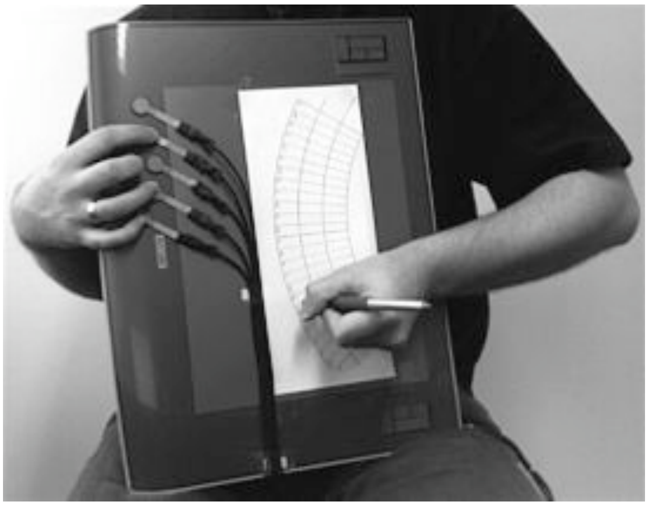

Figure 1: Representation of the "HandSketch" controller in use : Wacom™ tablet, with radial pen diagram, and 8 FSRs (Force Sensing Resistors).

2. PEN-BASED MUSICAL CONTROL

Graphic tablets, which are initially developed and sold to meet image professionals needs (designers, architects, editors, etc.), can today be considered as a common device in the computer music context. Moreover, in this field, we can observe unanimous use of Wacom™ products. As an example, in all existing NIME proceedings, more than 25 publications mentioned the use of a graphic tablet, and about 80 percents of them where Wacom™ devices.

Indeed, most of the models provide a large number of parameters, with high precision and low latency, structured around our intuitive writing abilities. These properties make Wacom™ controllers really good candidates to fit the Hunt and Kirk's *real-time multi-parametric control system* criteria [5][11]. Availability of Max/MSP and Pure Data externals, interfacing most of existing models, also contributes to this wide dissemination.

2.1 RealtimeCALM Control Model

Figure 2: RealtimeCALM gestural controller in use.

Last year, we presented new generators and mappings for expressive singing voice synthesis (RealtimeCALM) [3]. In that work, we proposed and demonstrated an example of gestural controller (*Instrument 2*), based on an A6 Wacom Graphire™ , in which pitch, intensity and timbre features (related to glottal flow parameters) were mapped to x, y and *pressure* stylus movements (cf. Figure 2). In spite of the simplicity of our mapping strategy (about 10 "one-to-one" or "one-to-many" connections), it appeared that realistic musical gestures could be reached (legato, vibrato, staccato, etc.) and that learning process (presently 1 year) evolved quite continuously until reaching an interesting level, allowing interpretation and improvisation [1]. Then tests have also been done with other kind of synthesis (especially acoustical instruments models) and our approach remained relevant. These "latent" skills are explained by the long background that we usually have in manipulation of pen, through writing techniques.

2.2 Pen-Based Pitched Sounds Control

As a matter of fact, we are focusing on the accurate control of pitched sounds[1] (singing voice in preceding studies [3][4]). In this field, litterature doesn't mention a lot of formalisms for pen-based continuous (*pitch*, *intensity*) gestures, as it exists of course for continuous pitch acoustical instruments, like violin [12], but also for some electrical devices, like theremin [8]. One of the most advanced formalization concerns the helicoid representation of notes in the Voicer [6], involving well know circularity in perception of frequency [9]. We also notice the Kyma [2] initiative, which developed a great framework for Wacom™ control of sound synthesis, but without formally considering (*pitch*, *intensity*) issues.

2.3 Solving Ergonomic Issues

Here we introduce a particular framework for expressive pen-based (*pitch*, *intensity*) musical gestures. This structure is much more based on ergonomic issues of playing and impact on sound synthesis, than on psychoacoustic representations. Our approach considers that natural pen movements are mainly forearm- and wrist-centered soft curves (cf. Figure 3), easier to perform than lines [3] or complete circles [6]. Then come fingers movements achieving details.

[1]This point of view doesn't exclude the use of our instrument for unpitched synthesis. We just mention that, in this work, devices, mappings and practices have been optimized for pitched sounds control.

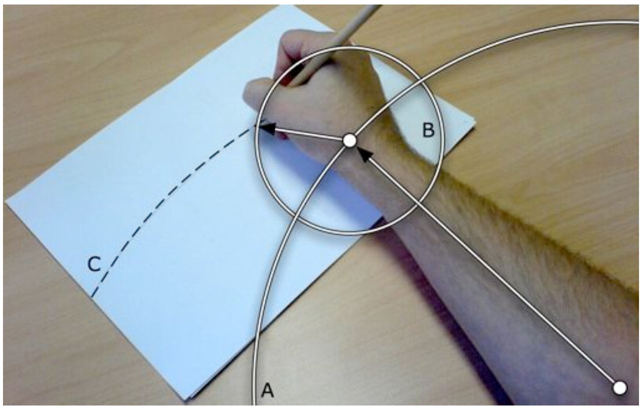

Figure 3: Pen drawing soft natural curve (C), as a combination of forearm- (A) and wrist-centered (B) movements.

Thus, we define a strategy where pitch information results from a transformation of (x, y) cartesian coordinates into polar coordinates, but where center of the circle position (x_C, y_C) is tweakable in order to fit forearm and wrist circular movements. Typically, this center will clearly be out of the drawing surface, close to tablet border, where the forearm is supported. This concept is part of the playing diagram illustrated in Figure 4. Equations are presented below :

$$R = \sqrt{(x - x_C)^2 + (y - y_C)^2}$$

$$\theta = arctan(\frac{y - y_C}{x - x_C})$$

with R and θ respectively the radius and the angle of the (x, y) point, measured in polar coordinates, with the center localized in (x_C, y_C), instead of (0, 0).

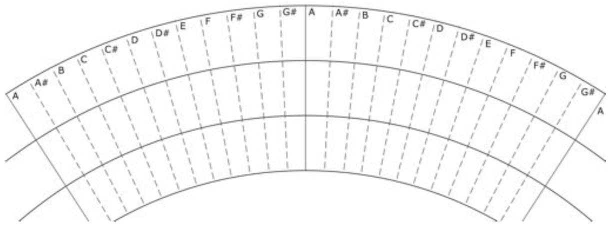

Figure 4: Diagram allowing performer to follow pitch angular representation, with tempered scale notes segmentation, and visibility for radius modifications.

2.4 Angle Mapping

As pitch control is now related to θ, angular information will be normalized and modified in order to lay out a range of notes (between 2 and 5 octaves) in which every semitone (in tempered scale) corresponds to the same angle. This representation is part of the typical playing diagram illustrated in Figure 4. This kind of conversion is obtained by following equations :

$$f_0 = f_{0_R} \times 2^{\frac{i}{12}}$$

$$i = N \times 12 \times \frac{\theta - \theta_B}{\theta_E - \theta_B}$$

where N is the number of octaves we want on the playing surface, θ_B is the leftmost angle visible on the playing surface, θ_E is the rightmost angle visible on the playing surface and f_{0_R} is the reference frequency corresponding to the θ_B position. A typical pitch modification on this diagram is illustrated in Figure 5.

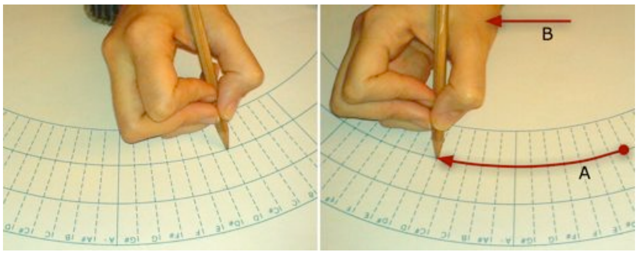

Figure 5: Demonstration of forearm/wrist movements (B) realizing a pitch modification (A).

2.5 Pressure Mapping

Concerning intensity mapping, we decided to keep the same approach as in RealtimeCALM control model [3]. Thus, sound intensity and stylus pressure are linked. It appears to be relevant, because based on the drawing metaphor, "making sounds" is related to "using pen", and pen is indeed used when pressed on the playing surface.

2.6 Radius Mapping and Interest in Finger-Based Gestures

A first way of highlighting radius (R) interest in this mapping, is to consider that some timbre features have to be controlled coherently with (*pitch*, *intensity*) gestures. A typical situation is singing synthesis control. Indeed, voice quality inflections often appear synchronously with pitch and intensity modifications, and combined control of these parameters effectively contribute to expressivity [4]. Thus, linking R with these spectral attributes justifies other kind of curves where radius is no more constant. Nevertheless, underlying forearm and wrist movements remain the same, and training just consists in integration of finger flexions. A typical mixed modification on the playing diagram is illustrated in Figure 6.

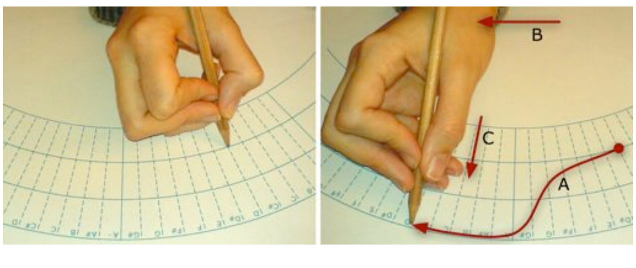

Figure 6: Demonstration of mixed θ and R modification (A) involving both forearm/wrist (B) and finger-based (C) movements.

Another interesting aspect of our layout concerns vibrato synthesis. Indeed, we know that oscillations do not concern only pitch, but also energy and several spectrum parameters [10]. In addition, it appeared that pen-based vibrato could easily be achieved by little circular movements around a fixed point. In this situation, both R and *pressure* are also involved in vibrato gesture description. Thus, this feature offers good opportunities to develop flexible multi-dimensional mappings around vibrato effects.

3. NON-PREFERED HAND ISSUE

Mapping strategies developed in section 2 proposed some ergonomic improvements, mainly in pitch and intensity manipulation. Performing on the diagram illustrated in Figure 4 makes it possible to learn and train to simple techniques (e.g. legato or vibrato), in order to reach an interesting level for interpretation and improvisation. Anyway, more advanced (*pitch*, *intensity*) structures, like arpeggios, trills, appoggiaturas, etc. are not possible. Moreover, even with the large number of parameters transmitted by the stylus[2], only slow timbre variations can be achieved. In summary, we observe that pen-based gestures have a inherent lack in controlling articultions of all kinds (pitch, intensity and timbre).

In this context, development of a controller for the non-prefered hand has to be done. Here we describe three main challenges that we propose to focus on. First, we need to be able to achieve quick and localized multi-finger sequences for (*pitch*, *intensity*) control, as used in many typical instrumental fingering techniques [12] (*direct control*). Then we consider the opportunity to use a priori music knowledge (e.g. tones, scales) in order to transform continuous (*pitch*, *intensity*) gestures into typical articulated sequences : arpeggios, long ascending and descending scales, etc (cf. "fine-tuning" non-linear pitch mapping in the Voicer [6]) (*modal control*). Finally, such a controller needs to be able to perform some timbre articulations, e.g. consonants (*spectral control*).

3.1 "5+3" Strategy

Considering preceding constraints, use of multiple pressure-sensing surfaces appears to be powerful. In this category, we can find several all-in-one controllers, such as Tactex MTC Express Pad™, Lemur™, or Z-tiles™. Anyway, we decided to develop an original "on-tablet" shape based on 8 independant FSRs from Infusion Systems™, for technical reasons : portability, unicity, price, latency, flexibility.

In this configuration, FSRs are separated into 2 groups. On the one hand, 3 sensors are aligned in order to define 3 thumb positions. These positions refer to our 3 main challenges (direct, modal and spectral controls). On the other hand, 5 sensors are aligned in order to achieve four fingers playing techniques. A major ergonomic issue of this configuration was to find a comfortable layout. As this problem could not be solved effectively with an horizontal tablet, it has been decided to flip the device vertically, in a position close to accordion playing (cf. Figure 1). Thus, group of 5 FSRs are placed on the front, and group of 3 FSRs on the rear part of the device. We observed that such a movement didn't affect writing abilities needed by the prefered hand. Details about positions are illustrated in Figure 7.

[2] Intuos™ models also send (x, y) pen tilt.

Figure 7: Demonstration of front and rear view of "5+3" diagram, with a typical hand position.

3.2 Direct Control

This technique has been developed to allow user to perform direct (*pitch*, *intensity*) modifications based on multi-finger playing techniques. It means that current pitch is build from the pen-based reference (related to θ), plus a deviation depending on adopted four fingers position. It can be used to implement e.g. "string-like" fingering sequences. A note pointed on the tablet corresponds to a "reference fret" on the virtual fretboard. Then pitch can be increased (3 semitones) or lowered (1 semitone) (cf. Figure 8).

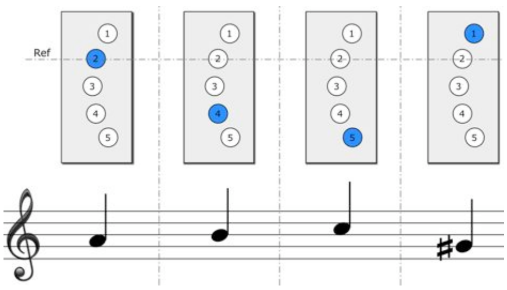

Figure 8: Illustration of a non-prefered hand "string-like" playing technique, with captor 2 as the "reference fret", corresponding to a *A4* pointed on the tablet.

3.3 Modal Control

This technique has been implemented to be able to perform large pen movements, with a structural control on harmonic contents. Thus, various finger configurations correspond to pitch and intensity non-linear mappings in a way arpeggios, defined scales or other note articulations can be achieved. Practically, pitch contour is flattened and intensity is modulated to sound louder around chosen notes. The amount of this effect is linked to FSRs pressure values.

3.4 Spectal Control

Concrete spectral techniques have not been implemented yet. Anyway, as movements on FSRs can reach a really interesting speed (presently a few events by second), this aspect has to be maintained as a major issue. Related e.g. to GRASSP project [7], the use of these kinds of fingering techniques could provide a powerful way to trigger phonetic contents in the context of real-time voice synthesis.

4. CONCLUSIONS AND PERSPECTIVES

In this paper, we have developed mapping strategies for both prefered and non-prefered hands. These investigations have been done considering some ergonomic aspects, in order to build a compact, intuitive, comfortable and expressive controller. Decisions were also made considering long-term underlying purposes, related to high-quality real-time articulated voice synthesis. Finally, it appears that features of this new controller clearly fit our expectations. Moreover, its abilities to let performer quickly create precise movements sequences, let us think about extending its possibilities, investigating e.g. polyphonic playing, more dedicated instrumental techniques (e.g. trumpet gestures) or complex mappings related to specific synthesis methods like physical modelling or concatenative synthesis.

5. ACKNOWLEDGMENTS

Authors want to thank FPMs University, TCTS Laboratory and SIMILAR European Network of Excellence for their logistics and financial support.

6. REFERENCES

[1] http ://www.dalessandro.be/files/nime06.mov.

[2] http ://www.symbolicsound.com.

[3] N. D'Alessandro, C. d'Alessandro, S. Le Beux, and B. Doval. Realtime calm synthesizer, new approaches in hands-controlled voice synthesis. In *Proceedings of NIME'06*, pages 266–271, 2006.

[4] N. D'Alessandro, B. Doval, S. L. Beux, P. Woodruff, and Y. Fabre. Ramcess : Realtime and accurate musical control of expression in singing synthesis. In *eNTERFACE'06 : 2nd SIMILAR Summer Workshop on Multimodal Interfaces*, pages 81–90, 2006.

[5] A. Hunt and R. Kirk. Mapping strategies for musical performance - trends in gestural control of music. *Trends in Gestural Control of Music*, pages 231–258, 2000.

[6] L. Kessous. Bi-manual mapping experimentation, with angular fundamental frequency control and sound color navigation. In *Proceedings of NIME'02*, 2002.

[7] B. Pritchard and S. Fels. Grassp : Gesturally-realized audio, speech and song performance. In *Proceedings of NIME'06*, pages 272–276, 2006.

[8] R. B. Sexton and D. Haussner. *Method for the Theremin Bk. I.* 1996.

[9] R. N. Shepard. Circularity in judgements of relative pitch. *Journal of Acoustical Society of America*, 36(12) :2346–2353, 1964.

[10] V. Verfaille, C. Guastavino, and P. Depalle. Perceptual evaluation of vibrato models. In *Proceedings of Conference on Interdisciplinary Musicology*, 2005.

[11] M. Wanderly and P. Depalle. Gestural control of sound synthesis. *Proceedings of the IEEE : Special Issue on Engineering and Music - Supervisory Control and Auditory Communication*, 92(4) :632–644, 2004.

[12] R. Woof. *Technique and Interpretation in Violin-Playing.* 2006.