Bimanuality in Alternate Musical Instruments

Loïc Kessous
CNRS-LMA
31 Ch Joseph Aiguier
13402 Marseille
0033 491 16 42 10
kessous@lma.cnrs-mrs.fr

Daniel Arfib
CNRS-LMA
31 Ch Joseph Aiguier
13402 Marseille
0033 491 16 42 10
arfib@lma.cnrs-mrs.fr

ABSTRACT
This paper presents a study of bimanual control applied to sound synthesis. This study deals with coordination, cooperation, and abilities of our hands in musical context. We describe examples of instruments made using subtractive synthesis, scanned synthesis in Max/MSP and commercial stand-alone software synthesizers via MIDI communication protocol. These instruments have been designed according to a multi-layer-mapping model, which provides modular design. They have been used in concerts and performance considerations are discussed too.

Keywords
Gesture control, mapping, alternate controllers, musical instruments.

1. INTRODUCTION
Since the beginning of human evolution, the hands of man were always privileged tools of expression. Be it for survival, communication, or artistic creation, few other elements can claim to play such a determining part in the relations and the interactions between the man and its social or material environment. Naturally, other elements such as voice, posture, glance, and facial expression play an important role in communication and creative activity. Musicians and researchers have explored some of these aspects [11] [18]. Bimanuality has always played an important part in conventional acoustic instruments be it in a coordinated or co-operative way. All these constituents must deserve all our attention when it is a question of defining new interfaces in the artistic domain, made of nuances and sensibilities. The possibilities brought by sound synthesis and digital effects have already changed our way of conceiving musical composition. Today, the composer in many different music currents is accustomed to not only write melodies, rhythms and harmonies, but also timbres and spectral evolutions. Today technology allows to play, to interpret, and possibly to improvise the sound itself. New musical gestures require new instrumental gestures or at least the reorganization of them to allow a real interpretation. In this article we will first introduce previous studies on bimanual skills from Human Computer Interaction. Then we will consider the concept of musical instrument and the mapping strategy that we use to design our instruments using alternate controllers. Three examples will be detailed. Finally we shall consider the performance situation.

2. HUMAN SKILLED BIMANUAL ACTION
Bimanual human behavior has been the subject of some interesting studies. Some efforts have been done to borrow tools for design and evaluation from Human Computer Interaction [1] [15]. Let us point out some interesting studies concerning this subject.

2.1 Bimanual coordination
In [7] Y.Guiard describes what he calls the kinematic chain model, a general model of skilled bimanual action (i.e a serial linkage of abstract motors). The KC model hypothesizes that the left and right hands make up a functional kinematic chain. This leads to three general rules:

- *Preferred-to-non-preferred reference:* The preferred hand performs its motion relative to the frame of reference set by the non-prefered hand.
- *Asymetric scales:* The preferred and non-preferred hands are involved in asymetric temporal-spatial scales of motion. The movements of the non-preferred are low frequency compared to the detailed work done by the right hand. The preferred hand acts efficiently at microscopic scales and the non-preferred hand at macroscopic scales.
- *Preferred hand precedence*: The non-preferred hand precedes the preferred hand; for example left hand positions the paper, then the right hand begins to write (for a right-hander).

Interaction between the two hands is a fact. It could be considered as a constraint in designing instruments, a benefit (if one knows it) for making them more efficient (when hands act on different tasks of a musical process), or the starting point for the design of a musical interface.

2.2 Cooperative bimanual action
In cooperative bimanual action, the two hands combine their action to achieve a common goal. Cooperative bimanual action is particularly effective to manipulate objects. These can be real [8] or virtual objects [2]. This kind of gesture is very interesting for performing electroacoustic music pieces. In this prospect, the analogy between visual or material objects and sound objects (as described in Schaeffer's theory [16]) is obvious. Using a multi-point tactile screen can be an interesting way to use cooperative movements.

3. MUSICAL INSTRUMENT
3.1 Acoustic musical instruments
The acoustic instruments have mechanical properties, which spontaneously offer to the performer force feedback, visual reference marks and multiple choices of selections. The force feedback and the visual and tactile reference marks allow an immersion in the playing and the regulation of the actions. Sense of touch, proprioception and visualisation provide different modalities of sensation. Multi-modality multiplies and diversifies sources of information, which allow a closer relation with the instrument.

3.2 Electronic musical instruments
3.2.1 Generalities
Generally, in electronic musical instruments, only the action of gestures is crucial to produce sound. Mechanical and

vibratory properties exist only if they are wished. Adding force feedback, visual feedback, visual and tactile reference marks must restore the epistemic function (the action of getting knowledge from the environment). Force feedback is not a standard programming feature on Macintosh, so we have concentred on visual and tactile reference marks which are easy to experiment (at very low cost). Some of our experiments seem to significantly improve the efficiency. On another side, electronic technology contributes to improve possibilities concerning the semiotic function (conveying information to the environment). Making the gestures symbolism understandable to computer is a good example of newly brought possibilities. In electronic musical instruments, the mapping plays a determining part because it makes possible to define the personality and the expressivity of the instrument. Measurement precision and interpretation of the ergotic function (the action of modifying and transforming the environment) are determininative points to allow the design of an efficient instrument.

3.2.2 *Choice of alternate controllers*

Existing controller can be classified as Instrument-like Controllers, Hybrid Controllers, or Alternate Controllers (see [19] [6] [12] [14] for examples and references). Our study is deliberately directed towards alternate controllers. What interests us is the possibility of defining an expressivity according to our wishes. Instrument-like controllers would impose in a large part the expressivity of their models. Hybrid controllers (acoustic instruments augmented by the addition of extra sensors) also do and require an expert control of the acoustic instruments. To start from more elementary data (position, pressure, flexion et cetera) makes it possible to conceive the expressivity on the level of the mapping without having to manage existing layers of mapping constrained by the physical configuration of the instrument. This approach seems to us most adequate to conceive the gestures corresponding to a sound and his evolution. A subdivision of Instrument-like Controllers is Instrument-inspired Controllers. This type of instruments is interesting in our musical context but one other reason to use Alternate Controllers relates to our interest for modular instruments. The postulate is to offer to the electronic performer instruments made up of couples of manual instruments. So the performer can choose the elements according to his needs.

3.3 Instrumental gesture typology

Various studies have defined different types of gesture. We use a typology close to the one defined in [3] and [20]:

- Excitation Gestures conveys the energy that will be found in the sonic result. It can be continuous, instantaneous or sustained. Static position that gives energy will be considered also as excitiation gesture.
- Modification Gestures modify the properties of the instrument but their energy do not participate directly in the sonic result. Such gestures can be divided into two groups: Parametric modification gesture - which continually changes a parameter (also called modulation gesture), and structural modification gesture - which modifies the structure of the object (instrument). Modulation gestures are more related to our study. Modifying a spectrum or applying vibrato will be considered as modulation gesture.
- Selection Gestures perform a choice among different but equivalent structures to be used during a performance. We will consider choice of musical note as a selection gesture.

4. MULTI-LAYER MAPPING CHAIN

Mapping can be used to define the personality and the expressivity of the instrument. It also allows defining an operating mode and a behavior in agreement with our sensori-motor system. We use three main steps between a gesture and a sound created by the synthesis model. The first mapping deals with the interpretation: it transforms gesture data into related-to-gesture-perception parameters, by means of a gesture extraction algorithm. This means that we transform measurements, quantitative data, into more qualitative information, closer to perception. By related-to-perception parameters, we mean parameters that make sense to our perception. This approach makes the implementation highly modular. This modularity is powerful to design instruments with a specific desired expressivity.

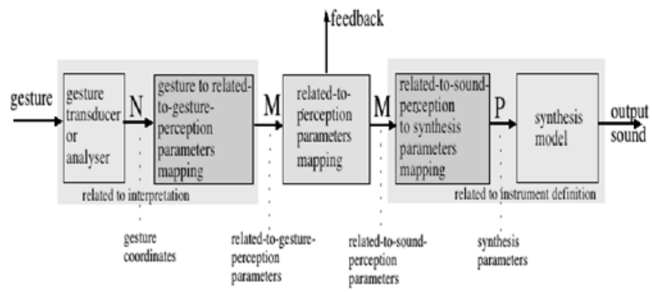

Figure 1. Three-Layer-Mapping chain

- In the first layer, gesture data are then transformed into related-to-gesture-perception data. Let's give an example take in literature. In [23], the parameters of three fingers are interpreted as a triangle (i.e. more significant parameters).
- The second layer transforms these data into related-to-sound-perception data. In [21], MIDI data from a wind controller is mapped to abstract parameters.
- Finally, a third mapping transforms related-to-sound-perception data into synthesis model data. In [22] pitch, loudness, and brightness are mapped to additive synthesis.

5. EXAMPLE 1 : THE VOICER

5.1 Description

The Voicer is a digital musical instrument producing vowel-like sounds with an expressivity depending upon the skills of the performer. It uses a bimanual control of subtractive synthesis with standard controllers. We have adapted a Wacom graphic tablet and a joystick to use them as an expressive instrument. Low-cost considerations, bimanual possibilities, and focus on making an efficient mapping have influenced the choice of two off-the-shelf controllers. Facilitating transmission to other (Mac addicted) musicians also encouraged us in this direction. This makes easier to evaluate the Voicer at expert level. It has been technically described in [10], so we shall shortly describe it here before going into bimanual considerations. The synthesis model consists of a sawtooth signal filtered by three second-order all-pole filters in cascade. This model can simulate a vowel singing voice. Here, we experiment with a mapping strategy that simultaneously allows melodic expressive control and vowel articulation.

5.2 Bimanuality in the Voicer

5.2.1 Preferred hand

To permit control within one octave and from one to the other, we divide the tablet's active space into 12 angular sectors where each sector corresponds to a semitone of the chromatic scale, but glides between successive notes are also permitted. Turning clockwise changes pitch from low to high. We can go from a note to its lower or higher octave by clicking the stylus lateral button up or down or incrementally by making a whole turn. Sound doesn't exist if there is no pressure (or MIDI aftertouch-like damping). The preferred hand carries out circular gestures when there is a continuous excitation (pressure), and gestures in quasi-straight line if not. In a first step we thought about controlling the voiced/non-voiced balance by the radius on the tablet, but we didn't do it because of too many influences from the other hand on this gestural parameters. We do not give up the prospect using it in another way, but this use was not adapted. In fact the influence of a trajectory of one hand to the trajectory of the other is an interesting point.

5.2.2 Non-preferred hand

The joystick controls the vowel articulation. X and Y data are mapped with perceptual attributes of sound described by Slawson [17]. In the case of vowel sounds, these attributes are closely related to tongue hump position and degree of constriction [5]. The non-preferred hand carries out gestures in straight line, but also, in an intuitive manner, circular arc. The non-preferred hand first selects the starting point of the articulation, then the other hand starts playing a note, then the non-preferred hand navigates in the vowel sound color space.

5.3 Mapping in the Voicer

5.3.1 First mapping layer

X and Y coordinates on the tablet are first converted into polar coordinates. Then the 360° angle is divided into 12 angular sectors. Variation within each sector is extracted separately and a transfer function is applied to allow glissando and vibrato. This gives rise to a modulation gesture [10]. Choosing a sector is a selection gesture, moving inside the sector is a modulation gesture. Clicking down or up on the stylus button is a pre-selection gesture. Time interval between 2 consecutive defined pressure data gives us an instantaneous gesture as in a percussive instrument (like a MIDI note velocity). Maintaining or varying the pressure allows to maintain or vary the energy: it is considered as a continuous exciter gesture, so the pressure data is scaled and we apply another transfer function to it in order to get the right perception. X and Y data from the joystick are just clipped and scaled.

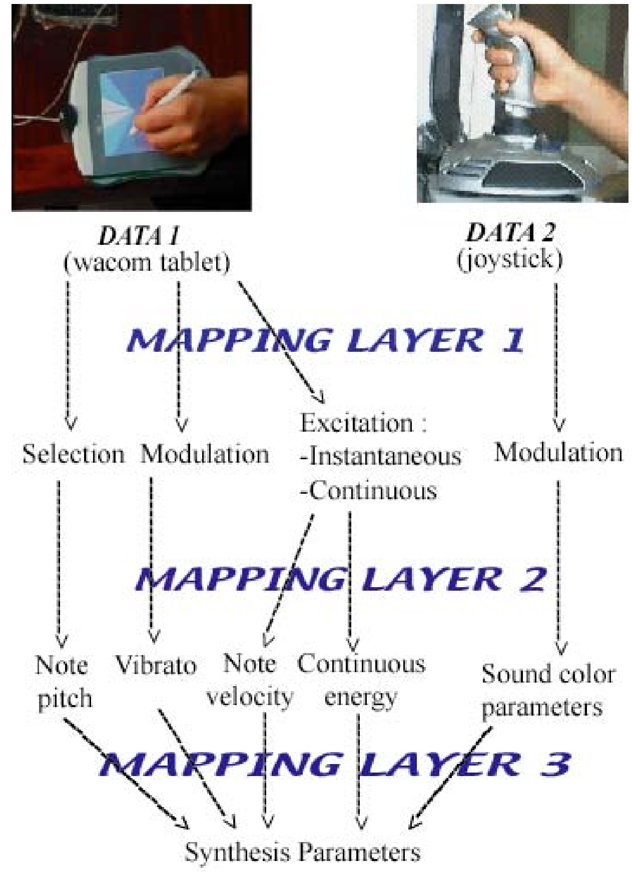

Figure 2. Three-Layer-Mapping chain for the Voicer

5.3.2 Second mapping layer

Joystick modulation gestures (or navigation) are assigned to a position into the plane of a 2D interpolator. Moving into this plane can be considered as 2 combined modulation gestures or as a navigation gesture. Selection gesture (selecting an angular sector) is assigned to a defined pitch. Here we define the register of the instrument, the way in which it is tuned (for a guitar, open string pitches are E, A, D, G, B, E), and possibly the tonal scale (1/2 tone, 1/4 tone, just scale, et cetera). Tablet modulation gesture (relative movement into a sector) is mapped to vibrato (like a MIDI pitch bend). Obviously, the tablet instantaneous excitation gesture is mapped to note velocity. Continuous excitation is mapped to continuous energy (in MIDI standard it could be volume, breath controller or others). By defining the interpolator parameters, navigation is mapped to sound color.

5.3.3 Third mapping layer

By using the interpolator, we interpolate the center frequencies of the formants (this mapping is clearly more conceptual), and sound colors are mapped to filter parameters (after conversion from frequency and radius). Note velocity and continuous energy are mapped to level and possibly (as we did in some experimentation) to parameters acting on source sound timbre (for example, when source sound is produced by Non Linear Distortion or FM). Note pitch and vibrato are mapped to fundamental frequency.

6. EXAMPLE 2 : SCANGLOVE

6.1 Description

The Scanglove is a bimanual instrument consisting of two different gloves equipped with sensors, controlling the parameters of a scanned synthesis [2]. In scanned synthesis, the shape of a simulated mechanic system is scanned at audio frequencies to produce sound. The scansynth~ object [4] generates a circular string (boundary conditions at the end of the string are transferred at the beginning) modeled with finite differences. We act on the initial shape and on the forces we apply. The synthesis meta-parameters that we have used are global damping, force gain and force extra-parameters (waveform "comb-like effect" suggested by Max Mathews and implemented by Jean-Michel Couturier). The non-preferred hand uses a 5dt Data glove [25]. The preferred hand uses a "home made" glove with FSR sensors (pressure and flexion) by way of an IcubeX (analog-to-midi converter).

Figure 3. Using The Scanglove on stage during "Glovy dub"

6.2 Bimanuality in the Scanglove

6.2.1 *Preferred hand (home made glove)*

This hand uses an explicit mapping velocity with pressure sensors which detect instantaneous excitation to trigger a note. Pressure is also mapped to continuous excitation. Two pressure sensors are positioned on first and second phalanx of the index and the thumb act on it. The upper sensor is used to trig notes defined by the other hands in a defined octave, the lower acts on the octave below. Two others flex sensors act on continuous parameters. The first one is placed on the middle finger and the second one on the little finger. Movement of the 4th finger is too dependent of the others fingers to be used. Middle finger and little finger are relatively more independent.

6.2.2 *Non-Preferred hand (5dt Data glove)*

For this hand we use the gesture's symbolism to do selection gesture. Using symbolism could be very helpful to drive complex harmonic structures while playing with the spectrum. A Multi Layer Perceptron (a Max external object available at [13], which can be trained in a real context) is used to map flex data from optic fibers of the 5dt glove to symbolic signs. The MLP external recognizes patterns of "Mimophony". The mimophony is a gestural code of a empty-hand symbolic sign representing a pitch note. It was recently used by a Contemporary Orchestra named "Allegro Barbaro" for conducting improvisation. From long time ago polyphonic singers from Corsica (a Mediterranean island near south of France) used it to communicate with each other while improvising.

6.3 Mapping in the Scanglove

6.3.1 *First mapping layer*

The 5 flexions from the 5dt data glove are mapped to symbolic signs. The pressures on the different sensors from the "home made" glove are mapped to continuous and instantaneous excitation gestures and selection gesture (simultaneous selection and excitation gestures could be interpreted as a decision gesture). The 2 flexions from the "home made" glove are mapped to modulation gestures.

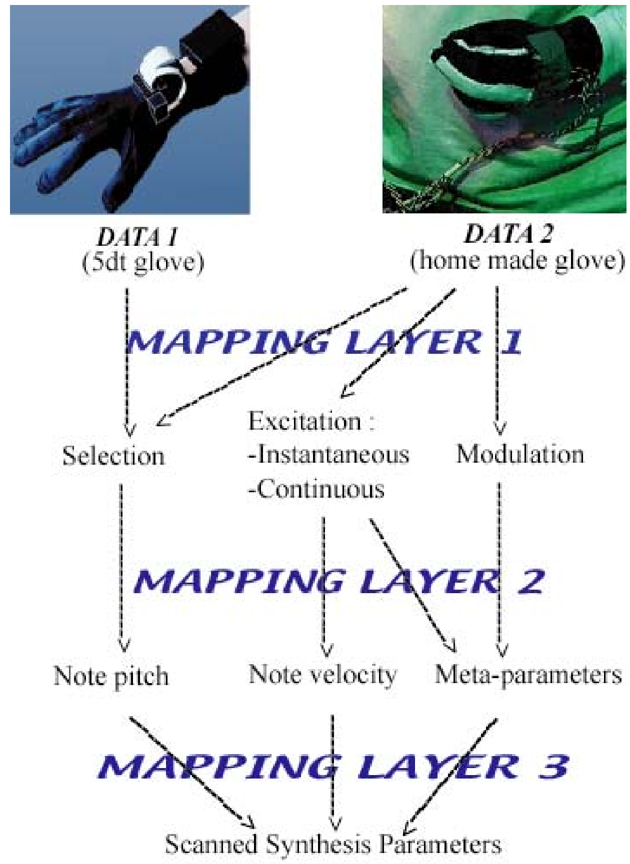

Figure 4. Three-Layer-Mapping chain for the Scanglove

6.3.2 *Second mapping layer*

Selection of sign is interpreted as selection among the 12 semitones of the chromatic scale, and selection of upper or lower pressure sensors is interpreted as choosing the octave: this is mapped to note pitch. Instantaneous excitation gesture is mapped to note velocity and continuous excitation gesture is mapped to force gain meta-parameters. The first modulation gesture is mapped to global damping, and the second modulation gesture is mapped to force extra-parameter (waveform "comb-like effect").

6.3.3 *Third mapping layer*

Note pitch is mapped to frequency, note velocity to initial shape properties, force gain and extra meta-parameters are mapped to low-level parameters.

7. EXAMPLE 3 : Voicer-like Controller

In this example we use the first two mapping layers of the Voicer with commercial stand-alone software synthesizers via the MIDI communication protocol (using IAC bus, or virtual MIDI port). Modularity provided by the three-layer-mapping chain makes it easy to use the expressivity (phrasing and articulation) of the Voicer to drive others synthesis models. In this case the last mapping layer can be done inside the software stand-alone synthesizer (most synthesizer are modular and can use GEN-like function). Some of them do not have strong-enough MIDI implementation to be driven. However, modulating pitch continuously (without triggering a new note) in a range of more than 24 semitones seems to be a real problem because the implementation is not available with enough precision. We have done this with the Voicer, going down to a sub audio frequency for which the glottal pulse is audible.

8. ON STAGE

We have used in concert bimanual instruments the design of which is based on the previously defined methodology. The band was composed of guitar, bass, drums, and a saxophone. Other electronic instruments (in particular, a wind controller mapped to a non-linear distortion synthesis model) were also used. The first challenge was to integrate a band and to adapt the instruments to the musical repertory. One does not have to permanently be in front of the computer to be able to communicate at any moment with the other musicians. Making gestures understandable to the audience is another challenge. Exaggerating gestures make them more comprehensible. A wireless technology could help in these two challenges, and probably more mechanical systems (which can also provide force feedback) could help for visual impact.

Figure 5. On stage with the band

9. CONCLUSION

We have shown in this article how to conceive, make and use digital musical instruments using a bimanual control. The principles exposed can help to design other instruments, and they give guidelines for their musical expressivity. The use in performance within a musical context illustrates the saying: "the proof of the cake is in the eating, not in the cooking".

10. ACKNOWLEDGMENTS

Our thanks to Jean-Claude Risset for comments and suggestions about this paper. Our thanks to Jean-Michel Couturier who has implemented the scansynth~ Max/MSP external object.

11. REFERENCES

[1] Beaudouin-Lafon M., "Moins d'interfaces pour plus d'interaction". In *Interfaces Homme-Machine et Creation Musicale*, Hermes Science Publishing, 1999.

[2] Buxton W. and Myers B., "A study in two-handed input", In *Proceedings of the 1986 Conference on Human Factors in Computing Systems* ACM CHI'86, 321-326, 1986. http://www.billbuxton.com/2hands.html

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

[4] Couturier J.M., "A Scanned Synthesis virtual instrument". *In proceedings of the 2002 Conference on New Instruments for Musical Expression* (NIME-02), pp. 176-178, Dublin, Ireland 2002.

http://www.mle.ie/nime/Proceedings/paper/couturier.pdf

[5] Flanagan, J.L., "Speech Analysis Synthesis and Perception". Springer-Verlag, New York, 1965.

[6] Fels, S. and Hinton, G., "Glove-TalkII: A neural network interface which maps gestures to parallel formant speech synthesizer controls*", IEEE Transactions on Neural Networks*, pp. 205-212, Vol 9, No. 1, 1998.

http://www.ece.ubc.ca/~ssfels/fpub.html

[7] Guiard, Y., "Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model". *Journal of Motor Behavior* 19 (1987), 486-517, 1987.

[8] Hinckley, K., Pausch, R., Proffitt, D., Patten, J., Kassell, N., "Cooperative Bimanual Action", In *Proceedings of the 1997 Conference on Human Factors in Computing Systems* ACM CHI'97, pp. 27-34, 1997.

http://research.microsoft.com/users/kenh/

[9] Hunt A., Wanderley M. and Kirk R., "Towards a Model for Instrumental Mapping in Expert Musical Interaction" *In proceedings of the 2000 International Computer Music Conference*, Berlin, Germany, 2000.

http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/Hunt_Towards.pdf

[10] Kessous L., "Bimanual mapping experimentation, with angular frequency control and sound color navigation". *In proceedings of the 2002 Conference on New Instruments for Musical Expression* (NIME-02), pp. 174-175, Dublin, Ireland, 2002. http://www.mle.ie/nime/Proceedings/paper/kessous.pdf

[11] Lyons M. and Tetsutani N., "Facing the Music: A Facial Action Controlled Musical Interface", In *Proceedings of the 2001 Conference on Human Factors in Computing Systems*, CHI'01 pp.309-310, Seattle, March 31 - April 5 2001.

http://www.mis.atr.co.jp/~mlyons/pub_pdf/Lyons_chi2001.pdf

[12] Merlier B., "A la conquête de l'espace", *Journées d'Informatique Musicale*, La Londe Les Maures, France, 1998.

http://tc2.free.fr/Espace/

[13] MLP, (short for "multi-layer perceptron"): Max Artificial Neural Network Object that implements simple backpropagation learning and forward-pass.

Available at:

http://cnmat.cnmat.berkeley.edu/Max/neural-net.html

[14] Modler, P., "Neural Networks for Mapping Gestures to Sound Synthesis". *In Trends in Gestural Control of Music*, M. Wanderley and M. Battier (eds), Ircam - Centre Pompidou, 2000.

[15] Orio N., Schnell N. and Wanderley M., "Input Devices for Musical Expression: Borrowing Tools from HCI" *Presented at the New Interfaces for Musical Expression Workshop* (NIME-01) during ACM CHI'01 - Seattle, USA, April 2001.

http://www.csl.sony.co.jp/person/poup/research/chi2000wshp/papers/orio.pdf

[16] Schaeffer P., "Traité des objets musicaux" Paris Seuil (eds), 1966.

[17] Slawson W., "Sound Color", Berkeley, University of California Press, 1985.

[18] Wanderley M., "Non-obvious Performer Gestures in Instrumental Music", in A. Braffort et al. (eds): *Gesture-Based Communication in Human-Computer Interaction*, Springer Verlag, 1999.

http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/GW99F.pdf

[19] Wanderley, M., "Performer-Instrument Interaction: Applications to Gestural Control of Music". PhD Thesis. Université Pierre et Marie Curie Paris VI, France, 2001.

http://www.ircam.fr/wanderle/Thesis/Thesis_comp.pdf

[20] Wanderley M., Viollet J-P, Isart F. and Rodet X., "On the Choice of Transducer Technologies for Specific Musical Functions". In *proceedings of the 2000 International Computer Music Conference*, Berlin, Germany, 2000.

http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/Wanderley_OntheChoiceFinal.pdf

[21] Wanderley M., Schnell N. & Rovan J., "Escher - Modeling and Performing Composed Instruments in Real-time", *IEEE Symposium on Systems, Man and Cybernetics* - SMC98 - San Diego, CA - USA, October, 14-17, 1998.

http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/SMC_FINN.pdf

[22] Wessel D., Drame C., and al. "Removing the Time Axis from Spectral Model Analysis-Based Additive Synthesis: Neural Networks versus Memory-Based Machine Learning", In *proceedings of the 1998 International Computer Music Conference*, Ann Arbor, Michigan, 1998.

http://cnmat.cnmat.berkeley.edu/ICMC98/papers-pdf/removing-time.pdf

[23] Wright M., Lee A., Momeni A., Freed A. "Managing Complexity with Explicit Mapping of Gestures to Sound Control with OSC", In *proceedings of the 2001 International Computer Music Conference*, Havana, Cuba, 2001.

http://cnmat.cnmat.berkeley.edu/ICMC2001/pdf/OSC-GestureMap.2.pdf

[24] Web site URL: 5dt glove. http://www.5dt.com/products/pdataglove5.html