A 3D GESTURAL CONTROLLER SYSTEM BASED ON EYECON MOTION CAPTURE SOFTWARE: A REVIEW OF 10 YEARS OF EXPERIENCE

Bertrand Merlier
Université Lumière Lyon 2
département Musique & Musicologie
18, quai Claude Bernard
69365 LYON Cedex 07
FRANCE
Bertrand.Merlier@univ-lyon2.fr

Frieder Weiss
Aurachweg 6
90449 Nürnberg
RFA

http://www.frieder-weiss.de/
info@frieder-weiss.de

ABSTRACT

This paper takes stock of a decade performing electroacoustic music thanks to a spatial gesture interface (3DGC), designed in 2002 for a complex instrumental control. 10 years later, motion capture technology has seriously improved, yet our device has not aged a bit.
The following lines recall its main features. Highlighted specificities appear to be perfectly suited to instrumental music, in a comprehensive and timeless way. The author shows that musical gesture control is not really a hardware problem, but rather a software problem i.e. making the good choices for gesture detection and analysis. The 3DGC gives many ideas on this topic.
Excerpt of 2004 paper introduction:
Both hands gestures are analyzed by means of a camera and by a motion capture software. Contrary to classical instrument interfaces, neither any mechanical organ nor any mark is visible! This concept offers important freedom of gestural control in any 3 dimensions of space. However it requires reconsidering the relation between cause and effect, i.e. between gesture and sound result. We shall present several specific features linked to lack of any visual mark or mechanical constraints. Then we shall underline the interest of programmability and simultaneous control of several parameters in only one simple gesture.
First use of 3DGC is to control generation, transformation and spatialization of live electroacoustic music. Many other artistic applications are imaginable in the multimedia area, such as live video pictures generation and transformation.

Keywords: HCI, Human Computer Interface, virtual instrument, computer-based musical instrument, gestural controller, motion capture software, electroacoustic music live performance, hand gesture analysis by means of a camera, Eyecon, MAX/MSP.

1. CONTEXT

1.1. Aim of our research

Our research fits into the context of creating innovative electronic music to be performed live. Composition and interpretation of “objets sonores” or “musique de timbre” frequently require a special job of “lutherie informatique”, in order to be able to simultaneously and effectively control all the numerous sound parameters.

1.2. Gestures types and uses

The present interface uses both hand gestures to accurately and effectively control[1] a sampler playing complex electroacoustic sounds (rich timbres whose heights are not necessarily defined) with multiple modulations possibilities (see pictures 1 and 3).

Using both hands motion tracking aims at replacing the synthesizer keyboard, leaving the musical keyboard paradigm[2] and the world of mechanics in order to obtain a new control interface with more degrees of freedom and greater wealth of play.

This virtual controller - detached from any mechanical constraint - has many degrees of freedom in three dimensions, such as: detection of "polyphonic" positions, accelerations, surface, form factors, symmetry, distance between both hands...

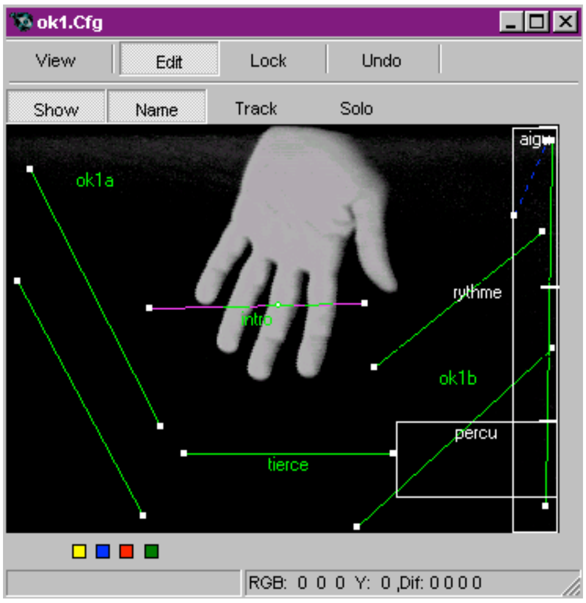
---

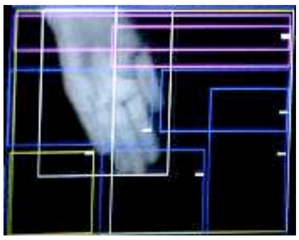
1 So called *online or direct manipulation gestures* opposed to *offline gestures*.
Source: http://en.wikipedia.org/wiki/Gesture_recognition

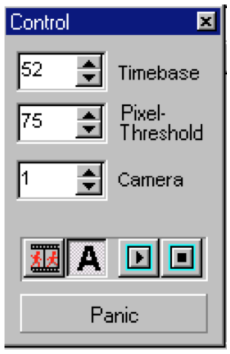
2 We have previously shown in various articles that the musical keyboard paradigm - focusing on simultaneous control of pitch and nuances - is far from being the better choice for an instrumental performance based on timbre :
Merlier Bertrand, « La main, le geste instrumental et la pensée créative ; CG3D, Contrôleur Gestuel Tridimensionnel », 2003 Actes des JIM 2003 (Journées d’informatique musicale), Montbéliard.
Merlier Bertrand, « La question des interfaces gestuelles dans les instruments électroniques du XX^{e} siècle : historique, état des lieux et présentation détaillée d’un Contrôleur Gestuel Tridimensionnel », non publié, 2008.

Picture 1: 3DGC overview.

The performer moves his hands into the camera field. The computer monitor offers a visual look back: the hand position in connection with the sensitive zones. If the sound result is obviously related to the gesture, the performer can rapidly consider forgetting to watch over the monitor.



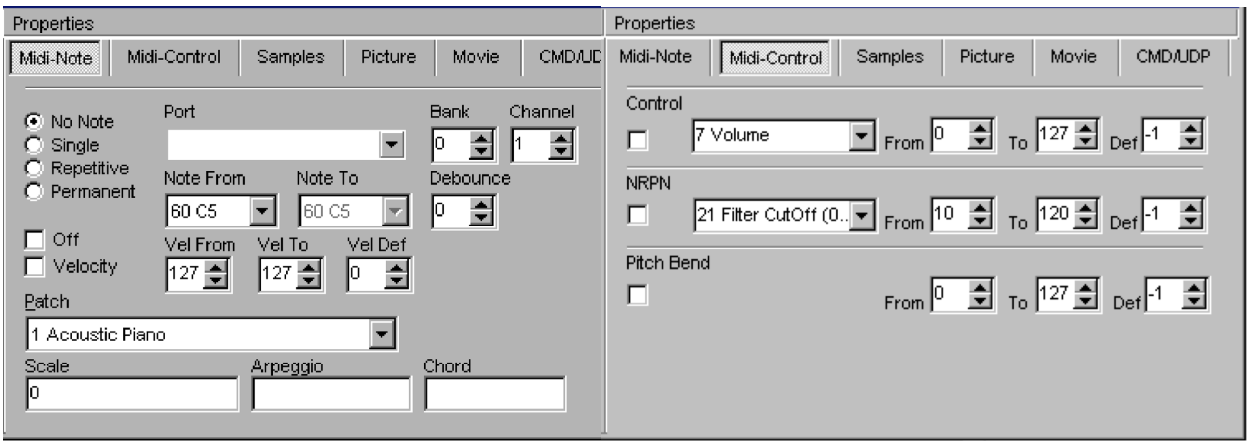




Picture 2: (above right) 3DGC adjustment panel for speed and grain detection

Picture 3: (on left and center) the *display* window

It first allows the building of the instrumental interface and then it is used as a visual control in the performance time.
In a single mouse click,
the detection mode can be changed.



Picture 4: the 2 MIDI settings panels (notes and controllers) associated to a detector

1.3. Birth of the device

In 2002, winner of an artist residency funded by the city of Dresden and the Art Festival CyNet, I designed a Tri-dimensional Gestures Controller (3DGC): a virtual instrument based on a sampler, driven by a both hands motion tracking system. This device has been used in dozens of concerts in Europe, in solo, in duets (music + video) or in trios. Project 4 Hands – a Multimedia instrumental duet – used two 3DGC for creating sound and live video. It was presented in September 2002 at Ars Electronica Festival (Linz Austria).

The 3DGC was the subject of several publications, including *Sound and Music Computing* at IRCAM - Paris in 2004.

Interests of the device are: affordability, simple standard and proven technology, small size, portability, reliability and accuracy in detecting the gesture...

1.4. 4 Hands

[4 Hands] is a multimedia instrumental duet using 2 3DGC. Two performers on stage are playing each their own invisible instrument. Jean-Marc DUCHENNE - video creator - performs live video sequences, cast on a main screen hanging at the rear of the stage and Bertrand MERLIER - composer - performs live electroacoustic music on 2 or 4 loudspeakers.

Four hands spin, caress, knock, scratch, stretch, describe complex arabesques, strange figures. The gestures of the two performers obviously control images and sounds, but no instrument is visible!

The pictures of the two performers' hands are also projected on two lateral screens. Soft lightings may also create hand shadows on the walls or on the ceiling.

The apparent simplicity of the equipment and of the stage design, the use of intimist lightings lead the audience to specially focus on the gestures themselves and on the main elements of the multimedia discourse.

The "4 Hands" project was presented in September 2002 at Ars Electronica Festival (Linz - Autriche), and then in several European electronic arts festivals.

Some presentation elements (music, pictures and video excerpts) can be seen at: http://tc2.free.fr/4hands/

2. 10 YEARS LATER…

2.1. A brief overview of the technology

As motion capture is concerned, we must distinguish two levels: the gesture capture device (input device) and the gesture analysis or interpretation algorithm.

2.1.1. New sensors

In 2012, the latest in 3D motion detection consumer technology are the Kinect[1], LeapMotion[2] or the system EyeSight[3]. These interfaces are essentially designed for video games. Many university or industrial research laboratories are also working on proprietary devices suited for more serious applications such as industrial process control, robotics, medicine, virtual reality[4].

These new consumer devices are more accurate, faster and less expensive than our 3DGC device designed in 2002. LeapMotion announces an accuracy of 0.01 mm at a cost of U.S. $ 70[5]. The Kinect is given to an accuracy of 1.3 mm. Most of the existing devices operate thanks to a infrared camera, in order to be insensitive to daylight or indoor lighting. Cameras standard speed of 25 to 30 Hz (25 to 30 frames per second) gives an analysis period of 33 or 40 ms.

2.1.2. Software for gesture analysis or interpretation

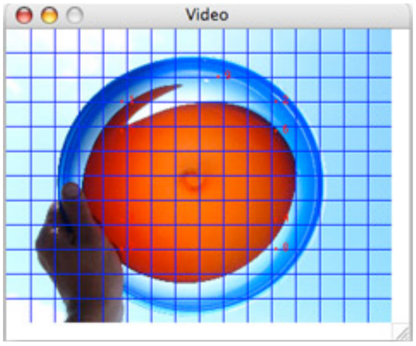
This is where the rub is when a composer does not have a computer lab devoted body and soul.
The problems are diverse:

a) As has already specified, main applications of these devices are video games: activity with little relevance to the musical gestures.
b) Regarding the Wii, Kinect or LeapMotion, one can find SDK (Software Development Kit)6 on the internet. Then the composer must forget his musical work and spend long hours of computer programming...
c) There are several software dedicated for gesture analysis (see pictures 6 and 7 and the note below[7]). Unfortunately, analysis algorithms provide data with little interest in music.

1 http://www.xbox.com/fr-FR/kinect/home or http://www.microsoft.com/en-us/kinectforwindows/
2 http://leapmotion.com/
3 http://www.eyesight-tech.com/
4 See for example: Proceedings of the International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999 or 2001, IEEE Conference Publications. http://ieeexplore.ieee.org/
Or else: Proceeedings of Workshop on Motion and Video Computing, 2009, IEEE Conference Publications.
Rehg, J.M. and Kanade, T., DigitEyes: vision-based hand tracking for human-computer interaction, Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 11-12 Nov 1994, p. 16 – 22.
5 http://live.leapmotion.com/about.html
6 Note the interesting OpenCV library (Open Computer Vision) for Python language: a free graphics library, originally developed by Intel, that specializes in realtime picture processing. Since 2010, the Willow Garage Robotics Society provides has been supporting for this Library. http://opencv.willowgarage.com/wiki/
7 IRISIS software developed by Collectif & Cie in Annecy - France: http://collectifetcie.free.fr/irisis.htm
Cyclops :Analyze and Track Live Video, developed by Eric Singer: http://cycling74.com/products/cyclops/

Reverse engineering has determined that the Kinect sensor outputs video at a frame rate of 30 Hz. The RGB video stream uses 8-bit VGA resolution (640 × 480 pixels) with a Bayer color filter, while the monochrome depth sensing video stream is in VGA resolution (640 × 480 pixels) with 11-bit depth, which provides 2,048 levels of sensitivity. The Kinect sensor has a practical ranging limit of 1.2–3.5 m (3.9–11 ft) distance when used with the Xbox software. The area required to play Kinect is roughly 6 m², although the sensor can maintain tracking through an extended range of approximately 0.7–6 m (2.3–20 ft). The sensor has an angular field of view of 57° horizontally and 43° vertically, while the motorized pivot is capable of tilting the sensor up to 27° either up or down. The horizontal field of the Kinect sensor at the minimum viewing distance of ~0.8 m (2.6 ft) is therefore ~87 cm (34 in), and the vertical field is ~63 cm (25 in), resulting in a resolution of just over 1.3 mm (0.051 in) per pixel.[1]

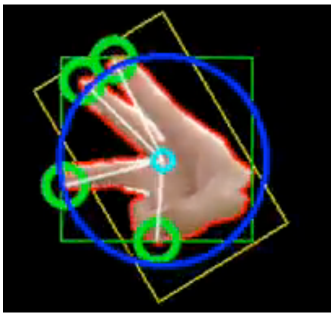
Picture 5: Kinect technical specifications

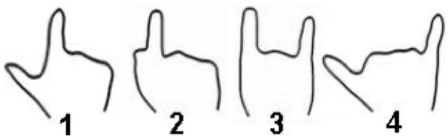


How Cyclops Works

Cyclops receives live video from a QuickTime input source and analyzes each frame of captured video in real-time. It divides the image area into a grid of rectangular zones and analyzes the zones for greyscale, threshold, difference (motion) and color. Cyclops allows you to specify the grid resolution, target zones for analysis and indicate the type of analysis to be performed in each zone. Cyclops outputs messages for each analyzed video frame that can be used to trigger any Max processes or control any patch parameters.

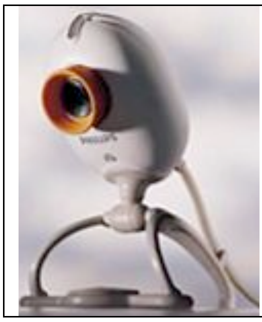
Picture 6: Cyclops software operation[2]





Picture 7: hand and fingers position detection

Min Yang Jung and Hyun Jae Kang, A real time Hand Motion/Gesture Detection for HCI, computer vision project, Johns Hopkins University, video available at: http://www.youtube.com/watch?v=sEGY6MyPqsY



	PCVC740K ToUcam Pro
Sensor	CCD
Pixels	640 (H) x 480 (V)
Still image resolution	1280 (H) x 960 (V)
Illumination	< 1 lux
Integrated lens	F2.0

Frame rate (frame per second) : up to 30 fps

Camera adjustment parameters:

Frame rate, contrast, brightness, gamma, saturation, colour on/off, mirror image, backlight compensation, white balance, exposure control

Picture 8: USB Philips ToUCam specifications

1 http://en.wikipedia.org/wiki/Kinect
2 http://cycling74.com/products/cyclops/

• The typical case is a 2D or 3D detection matrix (see picture 6): this mode of analysis focuses on orthonormal axes X, Y (and Z), it only has two (or three) degrees of freedom and happens to be much less virtuoso than any musical keyboard.

• Another analysis method consists in recognizing specific gestures or specific positions of the hand or the fingers (see picture 7) A typical application example is sign language recognition.

Ignorance of these issues along with a blind race towards modern technology can quickly lead to a significant loss of time and strong and regrettable organological regression going against the targets.

3. WHY NOT YIELD TO MODERNITY TEMPTATIONS?

Reasons are twofold:

Human and artistic factors:

a) a musical instrument requires a long training and a regular practice; organological instability usually means going backward in terms of virtuosity;

Four technological and organological factors:

b) detection grains control,

c) detection speed control,

d) detection primitives,

e) gesture-music mapping support (based on the MIDI standard).

After a brief presentation of our device, we will develop the final four points in the following paragraphs.

3.1. Brief reminder on 3DGC

Unlike acoustic instruments, the computer instrument 3DGC (Tri Dimensional Gesture Controler) is modular: one part capturing the gesture, the other part producing sound (or visual).

• Motion tracking is performed by EyeCon[1]. This software runs on PC. It requires a video capture board[2] and uses very little in terms of CPU resources. The computer display monitor returns pictures of the hands and the detection zones (see picture 3).

• In 2002, sounds production and processing was programmed in MAX/MSP. Today, the sound engine uses the Ableton Live software.

Hand gestures are filmed by a camera, analyzed by the software, then transmitted to sounds (or images) production means. Thus, they can trigger, control, transform any sound (or visual) material in real time (see Fig. 1, 3 and 4).

3.2. Detection grains

The instrumental performing space presents as an active volume[3] of about 1 m^3, which is huge compared to any acoustic instrument. The camera used is a cheap USB Philips ToUcam Pro. It provides a 640 x 480 pixels picture size.

Spatial precision: detection accuracy is about $1/150^{th}$ oh meters, i.e. about 1.5 mm. Compared to a piano key or a guitar fret, this precision is more than sufficient.

Space quantization: the software provides a detection grain control (called "pixels Threshold" on picture 2). For a grain of 16 pixels, the detection interval is 24 mm, i.e. the width of a piano key. However, realizing a 128 values MIDI controller requires a dimension of about 1.90 meters, well above the ambitus of the arm.

tempi /notes	𝅝	𝅗𝅥	♩	♪	𝅘𝅥𝅯	𝅘𝅥𝅰
♩ = 240	1000 ms	500 ms	250 ms	125 ms	62,5 ms	31,25 ms
♩ = 120	2000 ms	1000 ms	500 ms	250 ms	125 ms	62,5 ms
♩ = 60	4000 ms	2000 ms	1000 ms	500 ms	250 ms	125 ms
♩ = 30	8000 ms	4000 ms	2000 ms	1000 ms	500 ms	250 ms

Table 9: notes duration (in ms) versus tempo values

3.3. Speed detection

A standard camera operates at 25 or 30 frames per second, i.e. an interval from 40 to 33 ms per frame. Each image is divided into two interlaced fields of 20 or 16.6 ms each. So the detection interval is less than 20 ms.

Time accuracy: Table 9 compares this value with usual musical tempi. Even at high tempo, the detection interval is two times better than the sixteenth.

Temporal quantification: the software offers a detection speed control (called "Timebase" on picture 2); it allows somehow to fix the "intrinsic tempo of the gesture."

For a 125 ms interval, temporal quantification offers the following values: 125 ms, 250 ms, 375 ms, 500 ms ...
Or: 𝅘𝅥𝅯 , ♪ , ♪. , ♩ … at a ♩ = 120 tempo.

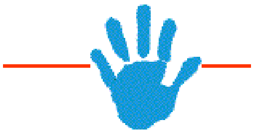
For a 166 ms interval, temporal quantification offers the following values: 166 ms, 333 ms, 499 ms, 666 ms ...
Or: 𝅘𝅥𝅯 , ♪ , ♪. , ♩ … at a ♩ = 120 tempo.

---

1 This commercial software (video motion tracking software) developed by Frieder Weiss is originally designed for dance and implementation of interactive multimedia environments: the dancer's movements control sound or multimedia elements generation. http://www.frieder-weiss.de/eyecon/index.html

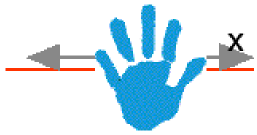
2 A Direct X compatible version can use any video devices (analog, DV, webcam...).

3 These parameters are determined by the camera specifications, the lens and the lighting mode. The volume is actually a truncated cone shape.

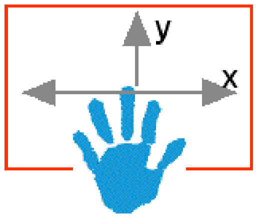


a) simple triggering: the hand crosses a line;

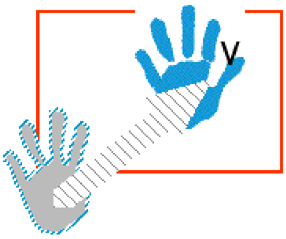
b) simple triggering: the hand goes into a detection box;



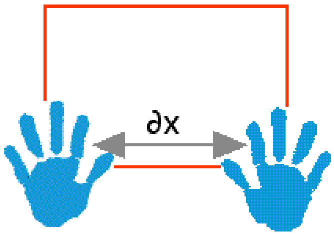
c) one dimension controller: the hand moves along a line;



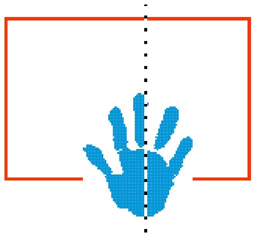
d) two dimension controller: the hand moves inside a box;



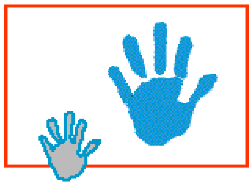
e) speed detection of the hand movement inside a box;



f) distance between both hands (tracking mode);

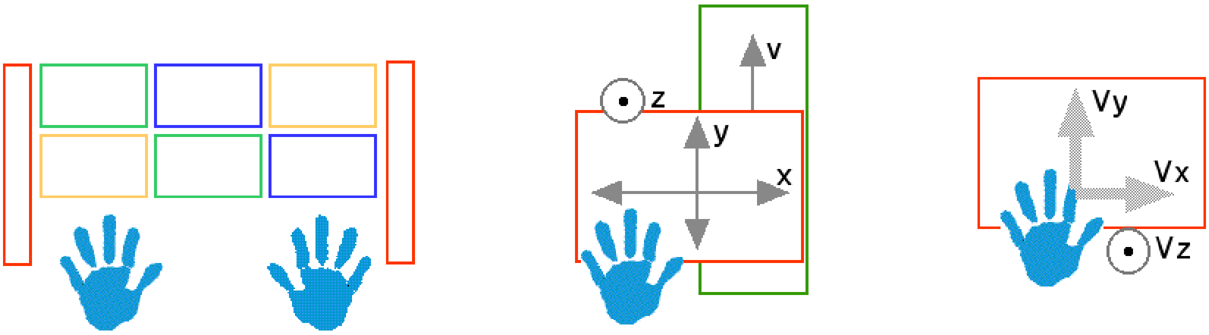


g) symmetry detection;



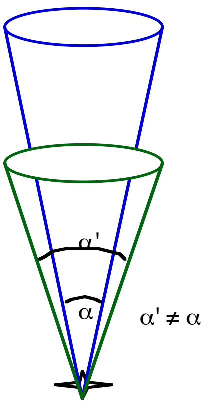
h) elevation detection (by measuring the lit surface);

Picture 10: elementary gestures detection with 3DGC



Picture 11: some complex instrumental gestures detections with 3DGC

a) events triggering by means of « on/off switches »;
b) multi-dimensional controllers (X Y Z V...) that can be superimposed or coupled together;
c) energy detection (by measuring the speed of the movements in the 3 dimensions X Y Z).



Picture 12: elevation measurement with the 3DGC

On the vertical axis (Z), the detection principle is completely different. The 3DGC measures the hand moving away from the camera by calculating the reduction of the lit surface (see figure 3h). In most cases, this way of proceeding is suitable.

3.4. Hand motion capture primitive functions

According to our composition needs, we used 6 types of instrumental gestures or movements:
1) a simple triggering, i.e. presence (or absence) of the hand (or the finger) in one space area (figures 10a and 10b);
2) X-Y position of the hand into a specific area (static mode) or its movement (dynamic mode) (figure 10c and 10d);
3) size of the lit surface; this parameter can correspond to the closure (clenched fist) or opening (flat position) of the hand, but also to the measurement of the distance of the hand from the camera (figure 10h and 12);
4) movement energy or speed inside a specific area allows to add a dynamic factor to the triggering (figure 10e);
5) trajectory following is possible, as well as relative distance between both hands, thanks to a tracking function (figure 10f);
6) last, it is possible to measure the left-right symmetry degree[1], as well as brightness, width, height or length of a shape (figure 10g);
Several occurrences of these six basic functions can of course be simultaneously used and freely combined together (figure 11), allowing really sophisticated gesture detection. We then have at our disposal an instrument with dynamic sensitive triggering sources and numerous controllers.

It is possible to superimpose a trigger and several controllers, so that one can trig an action with the clenched fist and then modulate it by opening the hand[2].

Combining several controllers together allows to dispose of a kind of a three-dimensional joystick (X-Y-Z) of about 1 m3 (figures 10d + 10h = figure 11b);

The speed (or energy) detection can be superimposed with any other detection area. With MIDI keyboards, we get used to "Note ON + velocity" coupling. This can easily be done with our 3DGC. A coupling such as "controllers + velocity" is much more original and opens the door to new exploration fields.

3.5. The gesture to music *mapping*

The 3DGC uses a fairly comprehensive MIDI implementation[3]: notes, continuous controllers, pitchbend on 16 channels and several ports) (see picture 4). The strongest supporters of modernity will certainly declare that this choice is a shame!

If the technology that supports MIDI standard is now largely outdated, however it appears that MIDI offers a reasonably accurate description of all the musical gestures in use. 30 years after its birth, that standard is universal: all the electronic music instruments adopted it and all types of computers now manage this standard via USB ports.

Our research is in no way questioning the purpose of musical instrumental gesture, but rather how to use these rich instrumental gestures to perform electronic music live on stage.

3.6. Conclusion

We presented the implementation of a gesture interface using (both) hands or forearm movements detection in three dimensions dedicated for production and control of musical or multimedia processes live on stage. The hand gestures are analyzed by a camera and a movement tracking software. This is a virtual instrument since 3DGC is achieved by computer programming and that there is no mechanical organ, nothing is visible.

3.6.1. The gesture spatiality

Acoustic instruments often have a preferred axis (X) generally corresponding to the heights control. The 3DGC provides an important gesture freedom in all the 3 dimensions of space, this in a volume of about 1 m^3.

3.6.2. New instrumental situations and new game modes

If the lack of any mechanical or visual feedback limits the melodic or rhythmic performance, however it favours a greater freedom of interpretation and a greater virtuosity in the work of timbre and dynamics, so far that one accepts to move away from the traditional instruments paradigms. The morphology and energy of the instrumental gesture are subtly linked to sounds.

3.7. Wealth of gestures detection and polyphony of the instrument

The 3DGC gesture interface simultaneously offers: many triggers and controllers, control possibilities of energy, coupled in a single gesture of the hand, allowing to establish causal complex relationships, therefore instrumentally interesting.

3.8. “Lutherie informatique” and programmability of computer interface

In our 3DGC, a few instructions can instantly change the number, arrangement, range and operation mode of any or all trigger and control zones. The software allows determining which instrumental gestures will be associated with what sound or image, what deformation or transformation. Finally, it is possible to program the relationship between instrumental gesture and the resulting sound, thus integrating “lutherie informatique” in the act of creation.

1 This is only a surface and symmetry factor measure but not a real pattern recognition.
2 In a MIDI keyboard, you need to use both hands (keyboard + wheel) or one hand plus one foot (pedal keyboard +) to obtain such a result.
3 For more complex applications, it is also possible to generate OSC codes.

3.9. Ten years later...

The present 3D gesture capture device designed in 2002 remains unequalled to this day.

If detection technology has progressed within 10 years, on the other hand musical gesture analysis software choice remains indigent or not adapted. The subject seems to be stagnating at a rather basic level, compared with oversized choice in the video games or virtual reality domains. While it is easy to kill a raging dragon in an intricate network of caves or to drive a racing car or a flying saucer through very crowded spaces, however it seems much more difficult to practice quieter activities such as performing music.

4. REFERENCES

4.1.1. Books

[1] Genevois Hugues et De Vivo Raphael (sous la dir. de), *Les nouveaux gestes de la musique*, Ed. Parenthèses, 1999.

[2] Halina Kwasnicka, Lakhmi C. Jain, *Innovations in Intelligent Image Analysis*, Springer, 2011 - 316 pages.

[3] Jianguo Zhang, Ling Shao, Lei Zhang, Graeme A. Jones, *Intelligent Video Event Analysis and Understanding*, Springer, 2011 - 250 pages.

[4] Miranda Eduardo Reck & Wanderley Marcello, New digital musical instruments: control and interaction beyond the keyboard (+ CD-ROM), Ed. CMDAS, 2006.

[5] Vinet Hugues et Delalande François (sous la dir. de), *Interfaces homme-machine et création musicale*, éd. Hermès, 1999.

[6] Wanderley Marcello and Orio Nicolas, Evaluation of input devices for musical expression: borrowing tools from HCI, CMJ, Vol. 26, N°3, Fall 2002.

[7] Wanderley Marcello et Battier Marc, *Trends in gestural control of music*, CD-ROM, IRCAM – Centre Georges Pompidou, 2000.

[8] *Workshop on Motion Capture for Music Performance*, October 30-31, 2006, McGill University, CIRMMT - Centre for Interdisciplinary Research in Music Media and Technology, New Music Building, 555, Sherbrooke Street West, Montréal, Québec, Canada.

4.1.2. Papers

[9] Bretzner Lars and Lindeberg Tony, « Use Your Hand as a 3-D Mouse... », *Proc. 5th European Conference on Computer Vision* (H. Burkhardt and B. Neumann, eds.), vol. 1406 of Lecture Notes in Computer Science, (Freiburg, Germany), p. 141--157, Springer Verlag, Berlin, June 1998.

[10] Duchenne Jean-Marc, Merlier Bertrand, « 4 Hands, an interactive live audio & video performance », in *Realtime and Presence, Trans Media Akademie*, Dresden, 2003.

[11] Erol A., Bebis G., Nicolescu M., Boyle RD., Twombly X., « Vision-based hand pose estimation: A review », *Computer Vision and Image Understanding* Volume 108, Issues 1-2, October–November 2007, Pages 52-73 Special Issue on Vision for Human-Computer Interaction.

[12] von Hardenberg Christian, Bérard François (2001), « Bare-hand human-computer interaction », *Proceedings of the 2001 workshop on Perceptive user interfaces. ACM International Conference Proceeding Series*. 15 archive. Orlando, Florida. p. 1–8.

[13] Laptev Ivan and Lindeberg Tony, « Tracking of Multi-state Hand Models Using Particle Filtering and a Hierarchy of Multi-scale Image Features », *Proceedings Scale-Space and Morphology in Computer Vision*, Volume 2106 of Springer Lecture Notes in Computer Science, pages 63-74, Vancouver, BC, 1999. ISBN 978-3-540-42317-1.

[14] Merlier Bertrand, « Three-Dimensional Gestural Controller based on Eyecon motion capture system », *Proceedings of SMC 2004 Conference (Sound & Music Computing)* - October, 20-21-22, IRCAM Paris.

[15] Merlier Bertrand, « New instruments for playing and spatializing electroacoustic music », *Symposium: Medienkunst - Verknüpfung der Sinne*, 27 Nov 98, COMTECart'98 - DRESDEN - RFA

[16] Merlier Bertrand, Les musiques interactives, vers de nouvelles relations entre l'interprète, le compositeur et l'ordinateur, *colloque « Musiques nouvelles »*, BORDEAUX, 1992 (publié en Mai 92 par le CEFEDEM Lyon)

[17] Pavlovich Vladimir, Sharma Rajeev & Huang Thomas S., « Visual interpretation of hand gestures for human-computer interaction: A review », *IEEE Trans. Pattern Analysis and Machine Intelligence*, July, 1997. Vol. 19(7), p. 677 -695.

[18] Quek, F., « Toward a vision-based hand gesture interface », *Proceedings of the Virtual Reality System Technology Conference*, p. 17-29, August 23–26, 1994, Singapore

[19] Rehg, J.M. and Kanade, T., DigitEyes: « vision-based hand tracking for human-computer

interaction », *Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 11-12 Nov 1994, p. 16 – 22.

[20] Schlomer Thomas, Poppinga Benjamin, Henze Niels, Boll Susanne, « Gesture Recognition with a Wii Controller », *Proceedings of the 2nd international Conference on Tangible and Embedded interaction,* 2008

[21] Sepehri Afshin, Yacoob Yaser, Davis Larry S., « Employing the Hand as an Interface Device », *Journal of Multimedia*, vol 1, number 2, p. 18-29.

[22] Stenger B, Thayananthan A, Torr PH, « Cipolla R: « Model-based hand tracking using a hierarchical Bayesian filter », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1372-84, Sep 2006.

[23] Wei Du, Hua Li, « Vision based gesture recognition system with single camera », *5th International Conference on Signal Processing Proceedings*, 2000.

Bertrand MERLIER

(Bertrand.Merlier@univ-lyon2.fr)

Electroacoustic and instrumental composer

Lecturer in Computer Music at the Lyon 2 University - France.

co-director of professional master MAAAV (Music Applied to Visual Arts)

For several years, he has been conducting researches on gestural interfaces in electroacoustic music, on interactive music and sound spatialization...

Author of "vocabulary of space" published in 2006 by Delatour Editions.

http://tc2.free.fr/merlier/