

Diseño de Interfaces de Control Basadas en Detección de Gestos.

Avance de investigación.

Fabián Sguiglia¹

Introducción

Los controladores diseñados para ejecutar instrumentos electrónico-digitales se inspiran en la mayoría de los casos en instrumentos acústicos tradicionales, complementándolos con dispositivos de entrada tales como perillas, pedales y sensores de proximidad. Este modelo permite construir interfaces útiles para la ejecución de música electroacústica pero, en algunos casos, resulta en la pérdida de una cierta inmediatez entre los movimientos de un intérprete y los sonidos que producen. En este contexto, la construcción de sistemas destinados a aumentar las posibilidades expresivas de instrumentos acústicos (hiperinstrumentos o instrumentos extendidos) y el desarrollo de instrumentos basados en control gestual (metainstrumentos) son tendencias de investigación ampliamente adoptadas. Del mismo modo, existe un extenso cuerpo bibliográfico relacionado con la parametrización (cuantificación y asignación de significado) de gestos y con directrices generales sobre su captación. La información respecto de los criterios selección y la descripción detallada de movimientos propios de distintas disciplinas son, por el contrario, relativamente escasas. En este sentido, y como se destaca en Rosa-Pujazón et al. (2013), los movimientos de los directores de ensambles musicales han sido sorprendentemente poco estudiados como modelo para interfaces de control gestual. Los diferentes roles en la ejecución de música contemporánea, particularmente, están acompañados de gestos específicos, cuyo estudio y documentación sería particularmente útil en el diseño de controladores de instrumentos electrónicos. En paralelo, la evolución de las

¹ fsguiglia@gmail.com | Licenciado en Composición con Medios Electroacústicos (UNQ). Becario de Formación en Docencia e Investigación Tipo B, Departamento de Ciencias Sociales, UNQ. Director: Lic. Mariano Cura. Programa Sistemas Temporales y Síntesis Espacial en el Arte Sonoro (UNQ). Director: Dr. Pablo Di Liscia.

tecnologías de sensado amplía incesantemente el espectro de los gestos que pueden ser captados y, por lo tanto, sus formas de aplicación. Esto implica una reelaboración permanentemente de los métodos orientados a utilizar el cuerpo de un intérprete como generador de señales de control. En particular, la renovación de sensores robustos y accesibles como Kinect y Leap Motion permite desarrollar software de fácil distribución.

El trabajo de investigación *Diseño de Interfaces de Control Basadas en Detección de Gestos* tiene como objetivos catalogar gestos utilizados en la dirección de ensambles y diseñar instrumentos e instalaciones interactivas que los apliquen en el control de parámetros musicales. Para ello se realizaron visitas a ensayos de grupos de improvisación y de intérpretes de música contemporánea. Mediante el método de observación participante, se elaboraron notas sobre los gestos y el lenguaje utilizado, la dinámica general del evento, así como (cuando correspondió) sobre la relación entre el director y los intérpretes. Estas sesiones fueron documentadas complementando videos RGB con capturas tomadas por la cámara de profundidad kinect y realizando entrevistas semiestructuradas con los directores e intérpretes. Los gestos observados en estas sesiones están siendo analizados para generar una clasificación a partir de descripciones pensadas específicamente para su detección a través de sensores. El objetivo de ésta comunicación es presentar el estado de dicho trabajo de investigación.

Recopilación y catalogación de gestos

Para realizar el muestreo de gestos utilizados por los ensambles entrevistados se diseñaron herramientas basadas en los sensores Kinect y Leap Motion. Kinect es una cámara de profundidad desarrollada por PrimeSense y luego adquirida por Microsoft para su consola de videojuegos Xbox 360. PrimeSense ha publicado, bajo la licencia LGPL, un framework llamado OpenNI, que permite acceder a la información de profundidad de Kinect y registrar las posiciones en tres dimensiones de los miembros de una persona ubicada frente al sensor. El sensor Leap Motion funciona de forma

análoga a Kinect, pero fue diseñado para obtener información detallada sobre la posición de las manos y sus articulaciones. En el marco del proyecto de investigación se escribió un programa en Processing destinado a obtener un registro combinado de los datos arrojados por ambos sensores. Estos datos fueron recopilados filmando directamente los gestos ejecutados en los ensayos. Cuando esto no fue posible, debido a las condiciones de iluminación o porque la colocación de los sensores resultaba demasiado invasiva, se solicitó a los directores que repitieran sus movimientos fuera del ensayo o se recurrió a imitar las señas utilizando como referencia las filmaciones.

En la comunicación gestual entre el director de un ensamble y sus integrantes entran en juego múltiples factores, que pueden incluir desde las expresiones faciales del director hasta la interpretación que hace de sus instrucciones algún miembro influyente del ensamble (Dineen 2011). Para dar cuenta de estas características, que no pueden ser capturadas por el sistema de sensores, se complementó la filmación de los ensayos con una cámara de alta definición con entrevistas semiestructuradas con los directores e intérpretes. De esta forma se pretendió también dar cuenta de las percepciones subjetivas de los actores, relacionar el movimiento con su significado y obtener información sobre el origen de las señas utilizadas.

Se ha llevado a cabo un primera análisis de los gestos capturados en base a las clasificaciones propuestas por Cadoz (1988 y 2000), y la extensión realizada por Godoy (2006), que combina esta clasificación con la terminología de Schaeffer (1966) para describir tres pares acción-sonido (*action-sound types*). Se encuentra en proceso una reelaboración en base al método de descripción conocido como Análisis de Movimientos de Laban y sus derivados (Laban, 1971 y 1974. Davis, 1970), en particular en los componentes “esfuerzo” y “forma”, que permiten dar cuenta de los cambios de dirección del cuerpo en los gestos, como estos fluyen y del esfuerzo involucrado. Del mismo modo, se considera como un antecedente relevante el análisis realizado por Bräm y Boyes (2001) sobre gestos expresivos realizados por directores de orquesta clásica con su mano no dominante. Allí se propone una clasificación de los gestos de acuerdo al dominio de origen de la metáfora en la que se basan, un concepto tomado de estudios de lingüística cognitiva sobre

lenguajes de señas utilizados por sordos.

Al momento se han entrevistado los siguientes ensambles y directores:

La Bomba de Tiempo, Percusión con Señas

La Bomba de Tiempo, fundada en 2006 por Santiago Vázquez, es un grupo de percusión que practica la improvisación guiada por señas. Los gestos utilizados por este ensamble fueron ideados por su fundador, complementando -según su propio relato- gestos tomados de la dirección orquestal clásica y del lenguaje de dirección desarrollado por Butch Morris con gestos originales. Santiago Vazquez ha elaborado un manual (Vazquez, 2013) donde describe los movimientos involucrados y los parámetros que controlan. Se han realizado entrevistas breves con Vazquez y se ha asistido como oyente a cursos donde presenta su lenguaje de señas. Del mismo modo, se han realizado entrevistas con Sebastián González, integrante activo de La Bomba de Tiempo.

Ensamble Argentino de Improvisadores / Esmeralda

El Ensamble Argentino de Improvisadores se formó en 2012 en el marco Primer Encuentro Argentino de Improvisación y Composición Musical. Este colectivo utiliza diversos lenguajes por señas para conducir improvisaciones en diversos lenguajes musicales. Para estudiar estos lenguajes se han realizado estudios preliminares a través del estudio de grabaciones en video de clases de Marcelo Moguilevsky y entrevistas en profundidad con Carlos Adriano Herrera, fagotista del Ensamble Argentino de Improvisadores y director del ensamble Esmeralda. Con el objetivo de comprender la dinámica de los encuentros de improvisación se participó como intérprete en el concierto Esmeralda III, bajo la dirección de Carlos Herrera.

Bandalup

Bandalup fue fundado en 2013 por un grupo de estudio conformado por alumnos de Santiago Vázquez y Marcelo Moguilevsky. Este ensamble utiliza principalmente las señas desarrolladas por Vázquez para la Bomba de Tiempo, agregando señas propias y tomando criterios para la improvisación de Marcelo Moguilevsky. En este sentido, si bien sus gestos son similares a los de los ensambles ya descritos, éstos son utilizados en ocasiones para indicar acciones diferentes. El análisis de las prácticas de este ensamble ha permitido realizar un estudio comparado sobre ambas técnicas. Para estudiar la dinámica de Bandalup se ha asistido a ensayos y presentaciones en vivo, y se han realizado entrevistas con su directora, María Fernández Cullen.

Sound Painting

Sound Painting es un lenguaje de señas inventado a mediados de la década de los 80 por el compositor Walter Thompson. Este movimiento ha ganado popularidad en los últimos años, trasladándose incluso a otras disciplinas, como las artes audiovisuales. Si bien no se han registrado ensayos que apliquen este juego de señas, se han realizado entrevistas con Mauro Sarachian, quien ha estudiado y utilizado este sistema. Para dar cuenta de las características puntuales de los gestos, se ha recurrido a los manuales y los videos publicados por Walter Thompson (2006).

Augusto Arias

Augusto Arias es compositor y director de orquesta, egresado de la Licenciatura en Composición con Medios Electroacústicos de la UNQ. Se han analizado, a partir de un video del evento, los gestos que utilizó en la dirección de su ópera Gruta de Ninfas, en el concierto realizado en el Teatro Argentino de La Plata en diciembre de 2014. Este video fue estudiado en compañía de

Arias, quién dio detalles sobre el objetivo de las señas, las replicó y las relacionó con las señas típicamente utilizadas en la dirección de Parciales de Gerard Grisey, obra sobre la que realizó un curso de dirección.

Clasificación de gestos

Los gestos realizados por los directores descritos en el punto anterior fueron filmados y posteriormente analizados con el apoyo de las entrevistas. A partir de esta información se realizó un listado que organiza las señas de acuerdo al sistema de sensores necesario para su captura (Tabla 1).

Fuente	Cantidad	Esqueleto Cuerpo	Esqueleto Mano	Combinación	Postura	Gesto Dinámico
Arias	15	6	4	5	6	9
Moguilevsky	31	16	10	5	14	17
Sound Painting	18	9	3	6	6	12
Vazquez	14	7	3	4	6	8
Bandalup	9	3	4	2	4	5
Total	87	41	24	22	36	51

Tabla 1. Cantidad de gestos de acuerdo al sistema de sensado necesario

En el marco de este trabajo se utilizaron dos sensores, Kinect y Leap Motion. Kinect permite registrar la posición de los miembros de un usuario en tres dimensiones, mientras Leap Motion otorga información similar para las articulaciones de las manos. De esta forma, para detectar una determinada seña puede ser necesario utilizar alguno de los sensores o una combinación de ambos. Otra característica determinante a la hora de diseñar estrategias de sensado, es si las señas se ejecutan a través de una postura estática (*posturas*) o a través de una combinación de movimientos (*gestos dinámicos*).

Del mismo modo se realizó un análisis funcional de los gestos relevados, partiendo de tres categorías básicas presentadas en Cadoz (2000):

- Gestos de excitación: Proveen la energía que eventualmente estará presente en el fenómeno percibido. Estos pueden ser instantáneos (el sonido comienza cuando el gesto termina) o continuos (el gesto y el sonido coexisten). En el caso de los gestos realizados por directores, se consideró gestos de excitación a aquellos que determinan el comienzo o el fin de un sonido.
- Gestos de modificación: Relacionados con la alteración de las propiedades de un instrumento sin que medie una transferencia considerable de energía. Pueden ser paramétricos (variación continua de un parámetro) o estructurales (modificación de categorías diferenciadas, por ejemplo agregado o remoción de partes de un instrumento).
- Gestos de selección: Elección de elementos similares de un instrumento, no provee energía ni modifica propiedades. Pueden ser secuenciales o paralelos.

Si bien éstas fueron pensadas para gestos realizados por instrumentistas, pueden ser traspoladas a aquellos realizados por un director. Para ello debe considerarse el efecto que las acciones estudiadas tienen en el sonido resultante, independientemente de una transferencia de energía². En este contexto, se consideraron gestos de excitación aquellos que determinan el comienzo o el fin de un sonido, gestos de modificación aquellos que modifiquen parámetros de la sonoridad general, y gestos de selección a aquellos orientados a seleccionar algún subconjunto dentro de un ensamble o a indicar a un instrumentista que que utilice un elemento específico de su instrumento. En tanto participan de la comunicación entre el director y el ensamble, algunos gestos que podrían considerarse *acompañantes* (ie. que no afectan la producción sonora en el caso de un instrumento, o que no pretenden transmitir una instrucción en el caso de un director) fueron incluidos dentro del análisis.

² Este mismo ejercicio se ha realizado en Cadoz (2000) donde se las aplican al análisis de los gestos utilizados en la ejecución de un controlador MIDI .

Fuente	Cantidad	Excitación	Selección	Modificación
Arias	15	3	3	9
Moguilevsky	31	7	6	18
Sound Painting	18	3	4	11
Vazquez	14	4	3	7
Bandalup	9	2	3	4
Total	87			

Tabla 2, cantidad de gestos por categoría funcional.

La categorización de gestos se encuentra aún en desarrollo, como parte del trabajo propuesto para la renovación de la Beca de Formación. Para el período 2015-2016 se plantearon los siguientes objetivos:

- Desarrollar métodos que permitan almacenar eficientemente la información sobre movimientos corporales arrojada por los sensores Kinect y Leap Motion.
- Evaluar e implementar estrategias de reconocimiento de gestos en función de su aplicabilidad en interfaces de control basadas en dichas tecnologías.
- Organizar la clasificación de gestos realizada en el período anterior, orientándola a su utilización a través de las estrategias de reconocimiento evaluadas.
- Evaluar e implementar estrategias que permitan relacionar dichos gestos con parámetros sonoros de forma coherente.
- Validar la utilidad de los resultados obtenidos a través de su utilización en *performances* e instalaciones interactivas.

Los gestos estudiados fueron incorporados en el desarrollo de diversas interfaces de control y metainstrumentos. A modo de ejemplo se presentará el instrumento *Cuenco*.

Cuenco

Cuenco, un metainstrumento en etapa de prototipo, fue pensado para controlar la aparición y modulación de masas granulares. El instrumento consiste de una serie de cuencos de arcilla -identificados con marcadores fiduciales- sobre una mesa, un sensor Leap Motion que capta gestos manuales, y una cámara RGBD Kinect que detecta la posición de cada cuenco y permite determinar si ha ingresado un objeto en alguno de ellos. Para iniciar la reproducción de un fragmento sonoro es necesario arrojar un objeto dentro de algún cuenco. Los sonidos resultantes pueden ser luego modificados en tiempo real a través de los gestos detectados por el Leap Motion. Las acciones detectadas por el instrumento pueden ser transmitidas vía OSC o MIDI a cualquier dispositivo.

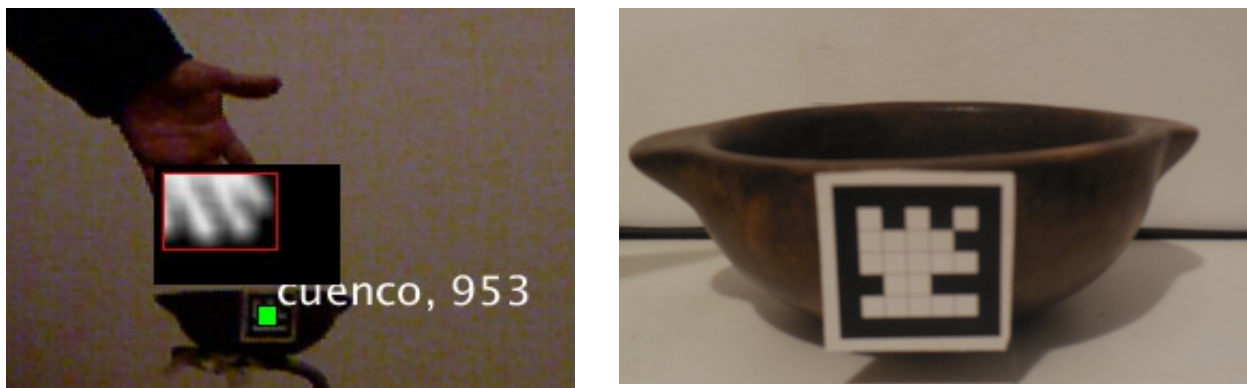


Figura 1. Izquierda detección de cuenco y blobs. Derecha, prototipo.

Para diseñar *Cuenco* se tomaron gestos utilizados por los distintos ensambles de improvisación estudiados. Estos gestos son detectados a través de las librerías *Leap Motion for Processing* y *\$1 Unistroke Recognizer* (Wobbrock et al. 2007). Al momento se han implementado los siguientes:

- Gestos de excitación
- **Arrojar un elemento dentro del cuenco:** La entrada de un objeto a cuenco determina el comienzo de un fragmento sonoro. El instrumento separa un objeto entrante de la mano del ejecutante a través de filtros por tamaño y evaluando el ángulo desde el que entra el

objeto y si este continúa por fuera del área de sensado.

- **Cierre de puño:** El cierre de la mano determina el final de un fragmento. Este gesto es utilizado por Augusto Arias.
- **Movimiento circular con la mano seguido de un cierre de puño:** Si el gesto de cierre de puño es precedido por un movimiento circular de la mano, el sonido cesa gradualmente. Este gesto es utilizado por Augusto Arias.
- Gestos de selección
 - **Posición del cuenco:** A través de la cámara Kinect puede detectarse la posición de cada cuenco. Esta información puede ser utilizada para elegir diferentes fuentes.
 - **Color del elemento arrojado:** La información arrojada por la cámara RGB de Kinect es utilizada para determinar si el objeto que ingresó en el cuenco es rojo, verde o azul. Esta información puede ser utilizada para elegir diferentes juegos de parámetros pregrabados.
 - **Indicaciones numéricas:** El ejecutante puede elegir diferentes estados a través de la cantidad de dedos que exhibe al sensor Leap Motion con su mano izquierda. Este gesto fue tomado del gesto “Memorias” utilizado por Bandalup y desarrollado en los videos de Moguilevsky. Para realizar este gesto, el director señala su cabeza y asigna un número con sus dedos al estado actual de la improvisación. La detección de la seña para indicar que un estado debe ser guardado aún no ha sido implementada.
- Gestos de modificación
 - **Movimiento de los dedos (“escribir en un teclado”):** La duración del gesto sobre el Leap Motion determina la densidad de las masas granulares. Este gesto es utilizado por Bandalup y por Augusto Arias como una indicación de textura. Fue extendido de tal forma que permita una modificación paramétrica para este instrumento.
 - **“Traer al frente”:** Este gesto, utilizado por Moguilevsky y Vázquez, indica a un instrumentista que debe proponer una idea. La seña se ejecuta con la palma hacia arriba, moviendo de forma reiterada los dedos hacia adelante y atrás. En Cuenco, este gesto se

realiza con la palma hacia abajo para hacerlo compatible con Leap Motion. Por el momento este gesto se utiliza para modificar el balance entre distintos cuencos.

- **Patrones rítmicos:** Esta seña, utilizada por Moguilevsky y Vázquez, permite al ejecutante seleccionar distintos patrones rítmicos eligiendo que dedos de su mano derecha muestra al sensor (Figura 2).



Figura 2. Indicación de patrones rítmicos.

Conclusiones

El trabajo de investigación *Diseño de Interfaces de Control Basadas en Detección de Gestos* tiene como objetivo generar una clasificación de gestos utilizados en performances artísticas en general y en la dirección de orquestas y ensambles de improvisación en particular, y aplicarlos al diseño de interfaces de control. La importancia de este tipo de interfaz no sólo reside en el enriquecimiento de la comunicación entre el artista y su instrumento, sino también en que impulsa una relación coherente entre el aspecto visual del gesto y su correlato sonoro: a fin de cuentas, casi toda la producción de la música adquiere sentido a través de la *performance*. Como parte de este trabajo se ha generado una base de datos de gestos, a partir de visitas a ensayos de ensambles y del

estudio de grabaciones de directores. Las señas analizadas están siendo organizadas en una clasificación y aplicadas en la construcción de diversas interfaces de control.

Bibliografía

BRÄM, T. y BRAEM P. *A Pilot Study of the Expressive Gestures Used by Classical Orchestra Conductors*, Journal of the Conductors Guild (2001): 14.

CADOZ, C. (1988), *Instrumental gesture and musical composition*. International computer music conference proceedings feedback.

CADOZ, C. y WANDERLEY, M. (2000), "*Gesture-music*". Trends in gestural control of music, nº12.

DAVIS, M. (1970) *Efort-shape analysis: Evaluation of its logic and consistency and its systematic use in research*. Four A daptations of Effort Theory in Research and Teaching, New York.

DINEEN, P. (2011), *Gestural Economies in Conducting*. New Perspectives on Music and Gesture 2011: 131. Ashgate Publishing, Ltd.

GODOY, R (2006). *Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus*. Organised Sound 11.02 (pp. 149-157).

SCHAEFFER, P. (1966) *Traité des objets musicaux*. Éditions Du Seuil, Paris.

LABAN R. (1971), *The Mastery of Movement*. Boston: Plays, Inc., Boston.

LABAN R. y LAWRENCE F. (1974) *Effort: Economy in Body Movement*. Plays, Inc., Boston.

VAZQUEZ, S. (2013) *Manual de Ritmo y Percusión con Señas*. Atlántida, Buenos Aires.

WOBBROK, J. (2007) "*Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes*." Proceedings of the 20th annual ACM symposium on User interface software and technology. ACM.