

Soporte Visual Asistivo para la Detección de Formas Matemáticas en el entorno

Visual assistance support for the detection of mathematical forms in the environment

Bruno Perelli S.

Universidad de Chile, Chile.

bperelli@uchilefau.cl

Pedro Soza R.

Universidad de Chile, Chile.

psoza@uchilefau.cl

Abstract

Nature exhibits forms of Vision without need for an eyeball. Brain adaptive abilities generates mental images allow to compensate for the visual stimulus captured through other senses such as touch and hearing. Technological extensions capable of mediating between the environment and the generation of sensory compensations, develop new forms of vision for the human being. This project proposes functional approaches capable of detecting visual elements in the environment that favor the construction of mathematical forms. The capture and processing of images mediated by the support, generate haptic sensory stimuli that performs the act of seeing without the need of eyes.

Keywords: Diseño; Computer Vision; Image Processing; Compensaciones Sensoriales.

Introducción

Cuando comúnmente se hace referencia a la visión, se establece una asociación directa con el acto de ver a través de un sistema de visión que implica la existencia de un globo ocular. La verdad es que la naturaleza presenta formas de visión diferentes del sistema de visión humano (Ings, 2008). Estos otros diseños de visión presentan diferentes enfoques a partir de la evolución particular de cada especie.

Así, el diseño de visión en especies como el Camarón Mantis (*Gonodactylus smithii*), el Topo Nariz Estrellada (*Condylura cristata*), el cual es capaz de ver a través del tacto, la Mosca común (*Musca Domestica*) y su ojo compuesto, la ecolocalización en los Murciélagos, entre otras formas de interpretar la información visual presente en el entorno, se muestra que en la naturaleza no necesariamente se requiere de la presencia de un globo ocular con las características técnicas del ojo humano.

En el caso de los humanos, además se asume que, de darse una deficiencia visual total o parcial, todo el sistema de visión se ve comprometido. Sin embargo, el globo ocular es uno de los múltiples componentes del sistema visual humano (Cattaneo y Vecchi, 2011; Ings, 2008). Por otra parte, técnicas desarrolladas para invidentes utilizan enfoques diferentes y complementarios para interactuar y ver el mundo. Tales son los casos de la utilización localización basado en el sonido ambiente, el olfato y el tacto (Cattaneo y Vecchi, 2011; Goldstein, 2009).

Adicionalmente, las capacidades de adaptación del cerebro se reflejan en la generación de imágenes mentales. Estas habilidades permiten derivar el estímulo visual capturado a otros sentidos como el tacto y la audición (Cattaneo y Vecchi

2011). Esta capacidad se conoce como Compensaciones Sensoriales.

Cabe entonces preguntarse: ¿Puede ver el hombre sin la necesidad de ojos? Y de ser así, ¿Cómo sería ese diseño de visión?

Para poder responder a ambas interrogantes, se debe advertir la diferencia entre lo natural, entendido como todo proceso que no implica la intervención del ser humano, y lo humano, lo cultural y por consiguiente lo diseñado.

En esa línea Hall y McLuhan (Hall, 1959; Leech y McLuhan, 1963; Zylinska, 2002) describen a la extensión como el medio capaz de liberar al cuerpo para el desarrollo de otras tareas, esa mediación tecnológica pretende dar respuesta a las incógnitas ante expuestas.

Este proyecto tiene por objetivo, diseñar una forma de visión compensada mediante la captura y procesamiento de formas matemáticas presentes en el entorno. Utilizando algoritmos de visión por ordenador y procesamiento de imágenes, el sistema de detección articula formas matemáticas básicas expresadas en puntos, líneas y ángulos. La función principal del prototipo es utilizar la información visual como transductor de datos para su posterior transformación en estímulos hápticos que construyan e indiquen el centro de un camino basado en la detección de formas matemáticas (Figura 1).

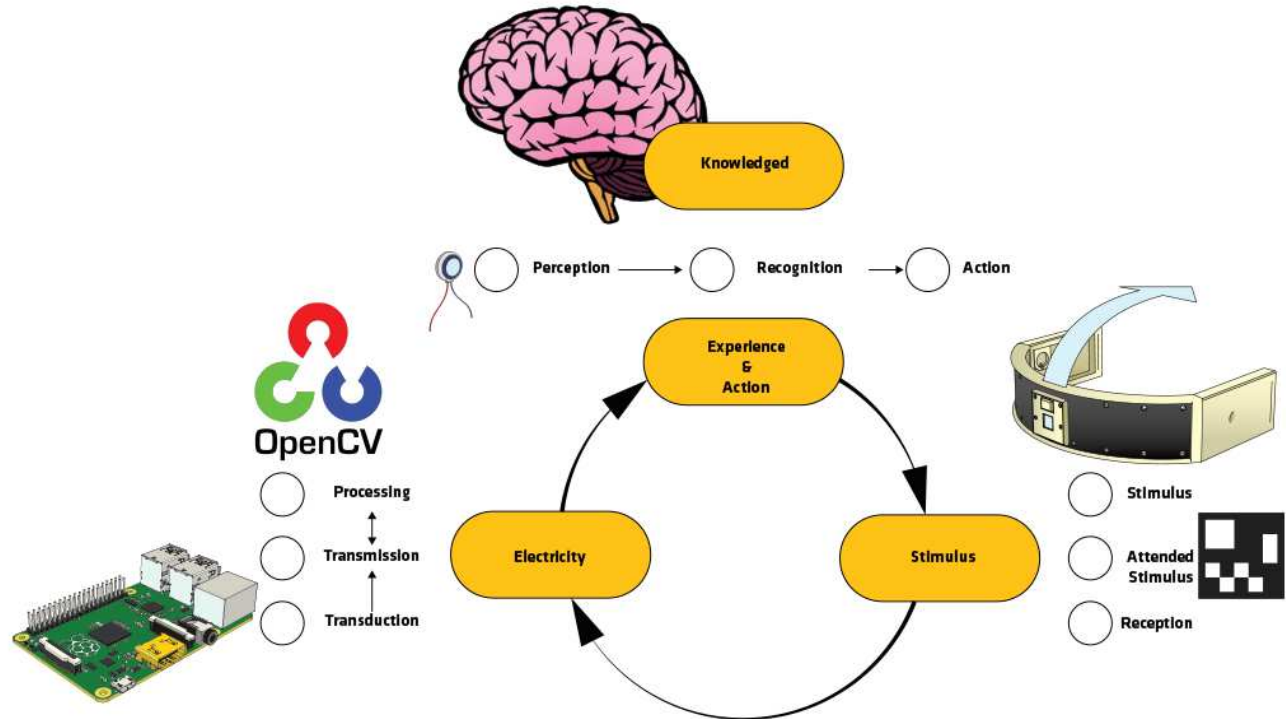


Figura 1: Medio de Visión Aumentada para Compensación sensorial Háptica basado en el modelo de proceso sensorial descrito por Goldstein. (Goldstein, 2009)

Metodología

El proceso inicial consiste de cuatro experiencias que exploran diferentes enfoques para el diseño de un sistema visual. La primera de estas experiencias se centró en la detección de color para señaléticas de tránsito dado sus formas, ubicación y color estandarizado.

Construcción de Software

Las directrices originales de la investigación, se centraban en tratar de generar la compensación sensorial basada en la detección del color dado que es la forma natural en la cual el humano es capaz de apreciar la realidad de forma más fidedigna. Sin embargo, la utilización de color mediado por computadores, implica limitantes técnicas no previstas como la estandarización de color por Hardware, Inconstancia del color y la cantidad de colores discriminados por el hardware y software de captura utilizado.

Aun así, las primeras experiencias de detección fueron bastante alentadoras. Se optó por operar con *OpenCV* dadas sus potentes bibliotecas utilizadas para la detección y procesamiento de imágenes. A partir de ello un script fue diseñado para la detección de muestras de color presentes en imágenes de la vía pública. Si bien la detección de señaléticas basado en su color demostró funcionar eficientemente, inmediatamente muestras de color pertenecientes a otros estímulos visuales presentes en la imagen plagaron la detección haciendo difícil de filtrar la data requerida (Figura 2).



Figura 2: La información de color no solo incluye información de las señaléticas de tránsito, además incluye muestras de otros estímulos visuales.

Una segunda iteración se centró en la detección las mismas señaléticas mediante la utilización de los algoritmos *SIFT* (Scale-Invariant Feature Transform), *SURF* (Speeded-Up Robust Features) y *ColorInRange*, función que detecta áreas de rangos de color presentes en imágenes o videos y los rastrea en tiempo real, los cuales a partir de una imagen referencial eran capaz de buscar la figura solicitada basado en sus colores y forma dentro de otras imágenes. Si bien esta iteración se caracterizó mucho más por el trabajo de laboratorio, presentó la gran limitante de que no todas las funciones podían operar en pistas de video y por lo tanto la captura de información en tiempo real tomaría aún más tiempo en la etapa de detección. Sin embargo, lo rescatable de esta implementación es que por primera vez durante el proyecto se

comienza a cuestionar el uso de color como estímulo visual y foco de la detección. Esto se da por la limitante de *SURF* y *SIFT* de buscar formas en escalas de grises y la limitante de *ColorInRange* que, si bien fue capaz de rastrear muestras de color expresadas en modo HSB (Hue, Saturation y Brightness), requiere de una constante calibración de iluminantes para que no se den inconstancias de color en la detección. Se centran los esfuerzos en generar un sistema visual capaz de detectar en tiempo real formas en desmedro de la información de color.

El tercer enfoque, se incluyó el análisis de los bordes (Edge) presentados en la imagen capturada por la cámara en las iteraciones anteriores. Este nuevo acercamiento se enfoca en el levantamiento de información visual relevante y detectable en desmedro de la detección y discriminación de color. De esta manera, los nuevos insumos de captura del entorno serán formas matemáticas tales como puntos, líneas y ángulos.

Para la implementación de este método, se utiliza un algoritmo diferente llamado Canny86. Este operador fue diseñado especialmente para capturar los bordes de la información visual presentada en las imágenes descartando la información de color. OpenCV indica la configuración de diferentes parámetros, tales como: Entrada de imagen (video o fotografía), mapa de borde de salida y el valor más pequeño entre el umbral1 y el umbral2 utilizados para el enlace de bordes. Cabe señalar que Canny requiere que la información visual venga en escala de grises para luego aplicar dos umbrales de desenfoque que le permiten mejorar el detalle de los bordes a detectar.

Para complementar la detección y por consiguiente la cantidad de formas matemáticas detectables a partir del borde de Canny, se usó una función llamada HoughLines. Esta función, dentro del algoritmo Canny, permite detectar líneas bien definidas en una imagen binaria. No obstante, independiente de su calibración, se detectó líneas que no correspondían a líneas regulares de trayectorias en espacios urbanos a partir del análisis de imágenes (Figura 3). La correcta calibración de los parámetros en ambos operadores fue necesaria. Para ello, se diseñó un pequeño programa que permitiera analizar imágenes de vía pública con la finalidad de encontrar la calibración correcta para Canny y encontrar las posibles líneas constantes en Houghs.

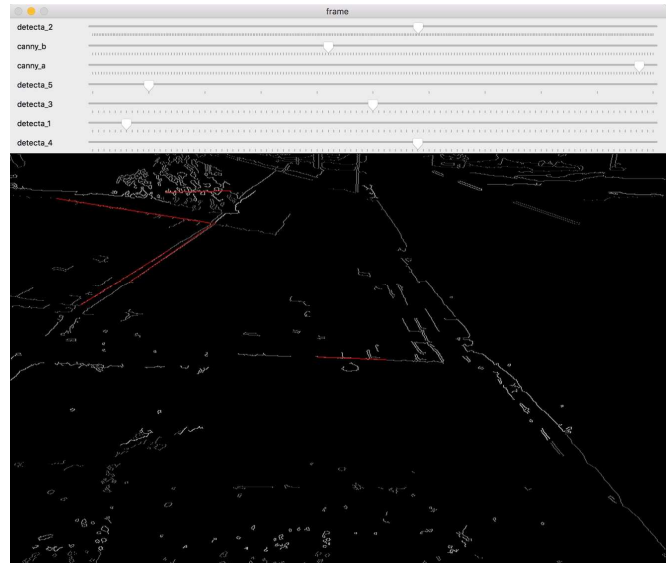


Figura 3: Script diseñado para el análisis de imágenes para la calibración de Canny86 y HoughLines.

Si bien se pudo construir líneas constantes a partir de la información de diversas imágenes de vías urbanas de desplazamiento pedestre, además de utilizar Canny en pistas de video, la calibración de HoughsLines, demostró ser demasiado sensible a cambios de parámetros detectando y dibujando líneas no deseadas. En la cuarta etapa del desarrollo de software, el enfoque sigue siendo el mismo con la salvedad de iterar nuevamente en el foco de detección. El fin de esta iteración se centra en la posibilidad de construir formas matemáticas a partir de información visual detectada a través de un dispositivo de mediación visual. Hasta ahora el gran problema a resolver es el foco o elemento a detectar. La multiplicidad de variantes que se da en una pista de video capturado en tiempo real genera información poco rastreable para su procesamiento y posterior construcción en respuestas hápticas, pero ¿Qué pasaría si se diseña también el estímulo visual a capturar?

La respuesta se encuentra el uso de Marcadores Fiduciaros o Quick Response (QR), este tipo de disposición permite albergar datos en una configuración binaria basada en la diferencia de matrices de color blanco (igual a 1) y negro (igual a 0). En técnicas de *Computer Vision* e *Image Processing*, los QR son utilizados además para establecer escalas en imágenes. Finalmente, su capacidad de albergar data binaria, permite que puedan integrar información para experiencias de Realidad Aumentada mediante el operador ArUco.

Este último método demostró ser mucho más eficiente dado que ArUco como operador recurre a Canny para la lectura correcta de los bordes de cada marcador. La experiencia final se enfoca en asignar un identificador numérico a cada marcador y mediante Realidad Aumentada establecer su punto centro en coordenadas (X,Y). Disponiendo dos sets de marcadores con el mismo Identificador numérico a cada extremo de un camino, los puntos centros de cada marcador serían detectados dibujando una línea que los uniera estableciendo el borde exterior izquierdo y derecho de cada camino. Finalmente, si se encontraban dos extremos, se

calcularían los puntos centros de la distancia entre ambos extremos dibujando la trayectoria que pasa justo por el centro del camino independiente de su configuración (Figura 4).

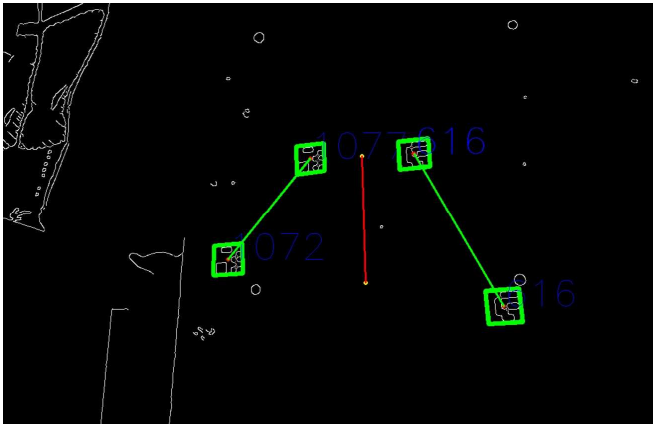


Figura 4: Prueba de Marcadores Fiduciales y Formas Matemáticas. El centro del recorrido se dibuja basado en los puntos medios de los extremos del recorrido.

Construcción de Hardware

Habiendo establecido la detección en el diseño de software, la elaboración de respuestas de hardware frente a la construcción de formas matemáticas se traduce en la construcción de compensaciones sensoriales. ¿Cuál es el o los sentidos ideales para compensar sensorialmente a la visión? Cattaneo y Vecchi (Cattaneo y Vecchi, 2011), aclaran que los sentidos más cercanos a la visión son los estímulos hápticos y auditivos. Sin embargo, tanto Jacobson (2013), como entrevistas con personas con ceguera adquirida demuestran la utilidad que los estímulos auditivos tienen en técnicas de desplazamiento pedestre. Por su parte, los estímulos hápticos utilizan las terminales nerviosas en la piel. El tacto en ese sentido, tiene muchos más canales de entrada (Grunwald, 2008).

Experiencias previas como Tactors de Paul Vach-y-Rita, utilizan la vibración de motores para la construcción de respuestas compensadas. Basado en dicha experiencia, y siendo el foco de detección los bordes y centro de un camino, se establecen tres tipos de respuestas a partir de la medición: un motor vibratorio para la detección del extremo izquierdo, uno para el derecho y finalmente un motor vibratorio se activará al detectar el centro del recorrido (Figura 5). La utilización de formas matemáticas permiten a su vez calcular el ángulo de la bisectriz, es así que si el valor de mi bisectriz (o ángulo del centro del camino) tiene un valor de $-0,11^\circ$, la respuesta vibratoria se localizará en el motor izquierdo. Si el valor es $0,11^\circ$ ocurrirá lo mismo al lado derecho. Si el rango de ángulos es de $-0,10^\circ$ a $0,10^\circ$ el usuario se encuentra en el centro del recorrido. De no haber detección, los motores se silencian (Figura 5).



Figura 5: Motores vibratorios DC para los estímulos hápticos de la izquierda, derecho y centro del recorrido.

Integración

Para la integración de Software y Hardware, la utilización de OpenCV y sus operadores y algoritmos demandó trabajar la programación con Python.

Dada su capacidad de albergar un sistema operativo capaz de trabajar con Python, su reducido tamaño y la capacidad de integrarse con otros elementos electrónicos, se decidió utilizar la placa Raspberry Pi 3 junto con la cámara de 5 Megapíxeles PiCamera.

Finalmente se montaron las piezas y cableado en unos lentes para ski y posteriormente se fabricaron las correspondientes carcasas para la cámara y la RP3 mediante impresión 3D (Figura 6).

En cuanto al diseño de los marcadores fiduciales se utilizaron inicialmente impresiones con la finalidad de realizar las primeras pruebas de desempeño.



Figura 6: Prototipo final.

Resultados

Para las pruebas preliminares se generaron dos recorridos. Uno de ellos fue una trayectoria en línea recta (Recorrido 1) mientras el otro presentaba una distribución irregular (Recorrido 2) debido a la distribución de códigos QR en el suelo. El monitoreo de actividad de los usuarios fue realizado de forma remota mediante VNC Viewer (software para acceso remoto mediante redes), dado que la RP3 es capaz de acceder a internet por su wifi integrado (Figura 7).



Figura 7: Pruebas de recorridos. Recorrido 2.

Desde el punto de vista de los resultados de la experiencia, la distribución de ambos recorridos requiere de un período de aclimatación por parte del usuario. Se cree que las condiciones lumínicas del entorno, la distancia de la cámara de los marcadores, además de la materialidad del QR requiere establecer condiciones óptimas para una detección más eficiente.

Sin embargo, cabe resaltar que la detección de marcadores en el *Recorrido 1* fue mucho más fluida que en el *Recorrido 2*. Por otra parte, pese a lo dificultad expuesta en ambos recorridos, fueron sorteados con un grado medio de dificultad.

Otro problema, esta vez de índole técnico tiene que ver con las fuentes de alimentación para la RP3. Una pequeña batería da una autonomía energética cercana a 20 minutos, situación que se debe mejorar en futuras iteraciones.

En iteraciones más recientes y consientes de la necesidad de profundizar en pruebas con usuarios se tomó la determinación de generar dos enfoques previos a nuevas experiencias con usuarios.

Discusión

Como experiencia inicial este primer acercamiento al diseño de formas alternativas de visión, pese a ser mediados por aspecto tecnológicos, ha sido todo un viaje. La reflexión inmediata que uno hace es evidenciar lo frágil que se han tornados los límites entre lo Natural y lo Human. Cómo el diseño e implementación de nuevos sistemas son capaces de aumentar, o suplir en otros casos, carencias y limitantes sensoriales en el ser humano tal como destacan Zylinska y Pepperell (Zylinska, 2002; Pepperell, 2003).

Frente al desarrollo del diseño de visión ha sido interesante el curso que fue tomando la temática desde la idea original hasta el diseño del prototipo final. Inicialmente el diseño de visión se enfocaba a la arista más conservadora del diseño: plantear un problema y hacer un objeto funcional para dar solución a dicha necesidad. En mi caso fue una búsqueda que partió por reconocer el color en la vía pública para luego dar una solución a los invidentes.

La proyección del diseño de visión a partir de los resultados preliminares ha generado variados caminos y aristas dada la confluencia de temas: Desde el punto de vista del diseño de marcadores referenciales se me abren caminos hacia temáticas como la integración de estos en el mobiliario urbano o el estudio de espectros visibles como la visión de elementos infrarrojos y ultravioleta. Creo que aún el prototipo es perfectible, debo reconocer que para esta experiencia elementos como la ergonomía han sido determinantes.

Dentro de esta investigación, se han considerado alternativas al uso de marcadores referenciales, como son luces LED infrarrojos o las geoetiquetas que utilizan algunas aplicaciones móviles basados en *google maps* como el ejemplo del popular *Pokemon Go*!

Por parte de los aspectos perfectibles, se advierte la necesidad de generar nuevos prototipos más funcionales capaces de leer con mayor eficiencia los marcadores (que como dije no considero que necesariamente deban ser QR). Aún faltan muchas otras pruebas con usuarios de forma tal de calibrar de mejor manera la detección y los tiempos de respuesta háptica.

Desde el punto de vista de la implementación de nuevas expresiones materiales con los marcadores, se han probado a la fecha diversos soportes como madera o acrílico, sin embargo, han demostrado generar sombras y brillos superficiales que Canny interpreta como nuevos bordes y por ende la detección y lectura del marcador fracasa o se ve

limitada a espacios. No se descarta optar por materiales como baldosas hidráulicas u otros soportes utilizados en rutas urbanas pedestres las cuales han demostrado ser aptos para su detección (Figura 8 y Figura 9).



Figura 8: Muestra de contraste blanco y negro en baldosas hidráulicas.

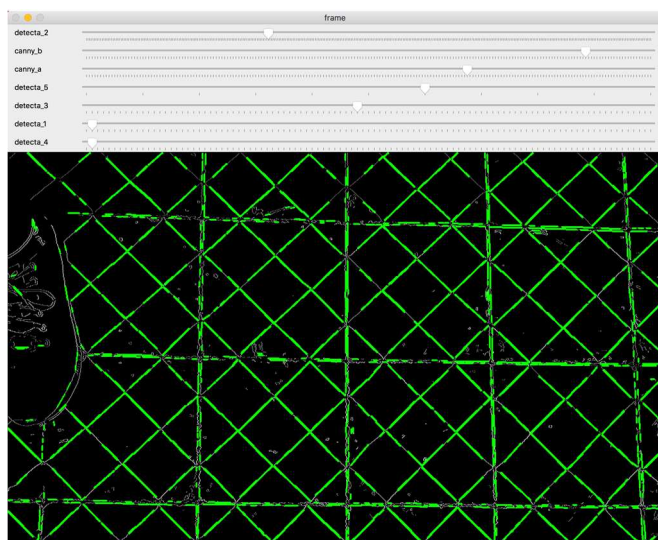


Figura 9: Análisis de líneas seguras en baldosas hidráulicas.

Es interesante reconocer el aporte del diseño a otros campos como la óptica, la visión por computador o el procesamiento de imágenes, el diseño no solo de los prototipos sino también de los sistemas con los cuales el sujeto se comunicará con el dispositivo, pasando por el desarrollo de marcadores referenciales que a su vez va mezclando variados aspectos que como diseñador siento que he tenido que saber adoptar y adaptar para llegar con este objeto al cual no quiero llamar prototipo final sino prototipo inicial.

Desde el punto de vista de futuras implementaciones, el mejoramiento en la manufactura de los marcadores y del software de detección y lectura sumado a la alimentación de energía permitirá entrar en etapas de pruebas con usuarios con la finalidad de evaluar la pertinencia de los estímulos hápticos como respuesta compensada a la discapacidad visual ya sea total o parcial.

Referencias

- Cattaneo, Z., & Vecchi, T. (2011). *Blind Vision: The Neuroscience of Visual Impairment*. Cambridge, Massachusetts: The MIT Press.
- Goldstein, E. B. (2009). *Sensation and Perception*. Retrieved from <http://www.amazon.com/dp/0495601497>
- Grunwald, M. (Ed.). (2008). *Human Haptic Perception: Basics and Applications*. Birkhäuser Verlag AG.
- Hall, E. T. (1959). *The Silent Language* (First Edit). GARDEN CITY, NEW YORK: DOUBLEDAY & COMPANY, INC.
- Ings, S. (2008). *A Natural History of Seeing: The Art and Science of Vision* (1 edition). W. W. Norton & Company.
- Jacobson, W. H. (2013). *Orientation and Mobility*. In R. Manduchi & S. Kurniawan (Eds.), *Assistive Technology for Blindness and Low Vision* (p. 29). Boca Raton, FL: CRC Press.
- Leech, G. N., & McLuhan, M. (1963). *The Gutenberg Galaxy: The Making of Typographic Man*. *The Modern Language Review*, 58(4), 542. <http://doi.org/10.2307/3719923>
- Pepperell, R. (2003). *The Posthuman Condition: Consciousness beyond the brain* (First Edit). Wiltshire: Cromwell Press.
- Zylinska, J. (2002). *Extending McLuhan into the New Media Age: An Introduction*. In J. Zylinska (Ed.), *The Cyborg Experiments: The Extensions of the Body in the Media Age* (First Edit). Continuum.