

Captura de movimiento del alumnado: una aproximación

F. Fernández Carrasco [1]; J. Botana Gómez; M. Sempere Tortosa [2]; C. González Gómez [1]; I. J. Navarro Soria [2]; D. Gil Méndez [3]

[1] *Departamento de Psicología Evolutiva y Didáctica*

[2] *Departamento de Ciencia de la Computación e Inteligencia Artificial*

[3] *Departamento de Tecnología Informática y Computación*

Universidad de Alicante

RESUMEN

Gracias a las nuevas tecnologías se ha conseguido contribuir a grandes avances en la ciencia y en la medicina, hacer intervenciones quirúrgicas menos invasivas, mejorar sustancialmente la calidad de vida de personas con necesidades especiales mediante aplicaciones específicas, ser mucho más precisos en el diagnóstico de trastornos y enfermedades e incluso la construcción de dispositivos que permitan a personas con algunas discapacidades físicas realizar acciones que de otra forma les hubieran resultado imposible. La tecnología no es incumbencia exclusiva de los tecnólogos. Éstos desarrollan las herramientas pero el verdadero sentido se adquiere con su uso, con su aplicación... Y precisamente esto es lo que nos ha llamado la atención a la hora de plantear este proyecto: cómo podemos aplicar las nuevas tecnologías para capturar y analizar el movimiento de los estudiantes y posteriormente buscar patrones de movimiento asociados a sus resultados académicos.

Palabras clave: captura de movimiento, kinect, tecnología, atención, patrones de movimiento.

1. INTRODUCCIÓN

Desde que comenzamos a preocuparnos e investigar el Trastorno por Déficit de Atención e Hiperactividad (TDAH), nos surgió la duda de si, en el ámbito universitario, podría tener alguna relación el rendimiento académico del alumnado con los tipos y cantidad de movimientos realizados durante el desarrollo de la clase.

Por ello vimos en el Programa de Redes de Investigación en Docencia Universitaria el marco ideal para empezar a trabajar en este proyecto, pues la confluencia de objetivos era clara.

Así, aunque este trabajo nace con el objetivo relacionar patrones de movimiento del alumnado con rendimiento académico, hemos querido ser realistas y empezar a enmarcar y acotar la investigación. Por ello, en este trabajo tratamos de hacer una primera aproximación a la captura de movimientos del alumnado en una situación natural de enseñanza-aprendizaje (en el aula), probando y examinando los diferentes instrumentos a nuestro alcance.

Tras hacer una primera revisión de la literatura existente, comprobamos que se ha relacionado el rendimiento académico del alumnado con múltiples variables y situaciones: con la atención plena o *mindfulness* (León, 2008), con el género y las redes de amistad (Requena, 1998), con el estrés (García et al., 2012), etc. Sin embargo, no hemos encontrado ni un solo estudio, investigación o artículo que intente relacionar el movimiento del alumnado con el rendimiento académico.

2. DESARROLLO DE LA CUESTIÓN PLANTEADA

2.1 Objetivos.

El propósito de este estudio es discriminar, analizar y describir la tecnología existente capaz de registrar movimiento para poder seleccionar una tecnología que cumpla dos requisitos:

- Que tenga validez ecológica. La validez ecológica es un tipo de validez externa que analiza el entorno de prueba y determina cuánto influye éste en el comportamiento.
- Que evite el reconocimiento de los sujetos participantes. En cumplimiento de la Ley Orgánica 1/1982 [1] sobre protección civil del derecho al honor, a la intimidad personal y familiar y a la propia imagen.

2.2. Método y proceso de investigación.

Una vez establecidos que definimos el objetivo, procedimos a la selección, prueba y elección del instrumento.

2.2.1. Selección de la tecnología e instrumento.

El registro de movimiento del alumnado podemos conseguirlo mediante tres tecnologías diferentes: vídeo, sistemas de captura de movimiento (MoCap) y sistemas capaces de generar mapas de profundidad.

2.2.2. Pruebas con los diferentes sistemas.

- Sistemas de vídeo.

El vídeo es la tecnología de la grabación, procesamiento, almacenamiento, transmisión y reconstrucción por medios electrónicos digitales o analógicos de una secuencia de imágenes que representan escenas en movimiento en dos dimensiones. Es la tecnología más accesible y sencilla para registrar movimiento. Además, para analizar el movimiento en un vídeo disponemos de un gran número de programas, como por ejemplo Tracker [2], Kynovea [3] o Dathfish [4].

Tras la realización de una serie de pruebas con los sistemas de vídeo, en relación con el objetivo de esta investigación, hemos encontrado una serie de inconvenientes que nos hacen descartar este tipo de instrumentos.

Frente a la ventaja de que es la tecnología más accesible, nos encontramos con que:

- Hay que realizar el análisis fotograma a fotograma, por lo que el tiempo y esfuerzo a invertir es inasumible.
 - No respeta el anonimato de los sujetos.
 - Únicamente permite trabajar en dos dimensiones.
- Captura de movimiento (MoCap).

La captura de movimiento (del inglés *Motion Capture* o *MoCap*) consiste en la medida y almacenamiento de las acciones directas de una persona real para analizarlas y reproducirlas inmediata o posteriormente (Thalmann et al., 2008). Esta técnica surgió en la industria del cine y los videojuegos para animar personajes diseñados por ordenador. En los últimos años se ha extendido su uso para fines deportivos (Hoffman et al., 2006; Leow et al., 2012) y médicos (Kanetaka et al., 2010; Sanmartín et al., 2012).

Existen diferentes sistemas de captura de movimiento que se clasifican según la tecnología que utilicen para realizar mediciones (ópticos, magnéticos, electromecánicos, inerciales, de fibra óptica, etc.). (Figura 1).

Figura 1. Diferentes sistemas de captura de movimiento



En la Universidad de Alicante se dispone de un sistema óptico de captura de movimiento al que hemos tenido acceso a través del Laboratorio de Creación de Contenidos Multimedia [5].

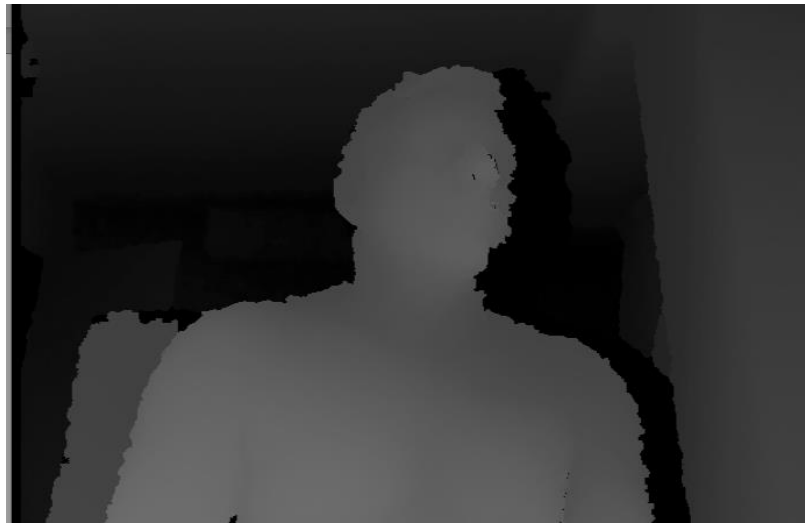
Los sistemas de captura de movimiento ópticos utilizan marcadores como material retroreflectivo para reflejar la luz que es generada cerca de las lentes de las cámaras. Monitorizando las posiciones de los marcadores es posible obtener las localizaciones que corresponden con las diferentes articulaciones. Para esto, se genera un esqueleto virtual sobre el cual se aplican en tiempo real los movimientos que realiza el actor.

Este sistema lo descartamos porque no se ajusta a los objetivos perseguidos.

- Sistemas de detección de mapas de profundidad.

Un mapa de profundidad es una representación en dos dimensiones de una escena en tres dimensiones. Generalmente consiste en una imagen que contiene información relacionada con la distancia a la que se encuentran los diferentes objetos del campo de visión. Para representar la información tridimensional se utiliza una escala de colores. Por ejemplo, para una escala de grises, a cada píxel de la imagen se le asigna un nivel en esa escala, representando en blanco el punto más cercano al dispositivo y en negro los puntos más alejados.

Figura 2. Ejemplo de mapa de profundidad



La figura 2 muestra un ejemplo de mapa de profundidad. En ella podemos observar cómo el busto de una persona es el objeto más cercano a la cámara puesto que está en un nivel de gris más claro. El respaldo de la silla se observa que se encuentra un poco más alejado ya que está en un gris algo más oscuro. Finalmente, el fondo de la habitación se encuentra en color negro.

Una de las características de esta tecnología es que permite detectar un cuerpo humano. Para esto, generalmente, cuando se detecta un cuerpo, se representa mediante un esqueleto. Este esqueleto consiste en un número de segmentos (como antebrazo, mano o pie), los cuales se conectan entre ellos por articulaciones (como codo, muñeca o tobillo) (Thalmann et al., 2008).

Esta tecnología, a priori, cumple los objetivos que nos habíamos fijado inicialmente, pues, por una parte, la podemos introducir fácilmente en el aula y, por otra, podemos recoger datos de forma totalmente anónima.

2.2.3. Elección de la tecnología.

Una vez descritas las características de los diferentes tipos de tecnología existentes, la que más se ajusta a nuestros objetivos son los sistemas de detección de mapas de profundidad. Tras hacer una búsqueda exhaustivo, en este apartado se describirán algunos de los dispositivos que utilizan dicha tecnología. Dada la gran variedad de dispositivos existentes, nos detendremos únicamente en aquellos que hemos considerado más relevantes.

- MESA IMAGING SWISSRANGER [6]

Dispositivo que utiliza un sistema de imagen óptico TOF (*Time-of-flight*) que proporciona datos de distancia en tiempo real mediante la señal de luz entre el dispositivo y el

sujeto. Las mediciones usando el principio de TOF se realizan a partir de 5 metros. Se pueden desarrollar aplicaciones utilizando librerías como IISU SDK de SoftKinetic, OpenCV (lenguajes de programación C y C++), PCL (Point Cloud Library), incluso Matlab. Estos dispositivos se utilizan en laboratorios académicos, pero su alto coste de adquisición es un inconveniente a tener en cuenta (Figuras 3 y 4).

Figura 3. SR4000



Figura 4. SR4500



- PMD CAMCUBE [7]

Se trata de un sistema óptico de última generación desarrollado por PMD. Dispositivo de alta resolución y ratio de imágenes (FPS). Tiene un diseño modular y realiza mediciones usando el principio de TOF en un rango de 7.5 metros extensible con un campo de visión de entre 40 y 60 grados según el tipo de lente que se utilice. El dispositivo es compatible con librerías de programación tales como Omek Beckon SDK, OpenCV, PCL y PMDSDK2 (librería proporcionada por el fabricante del dispositivo). Aunque en su descripción del dispositivo PMD indica que es una cámara de profundidad de bajo coste, no se ha encontrado su precio de mercado ni la posibilidad de comprarlo, pese a ello se han encontrado reseñas en las que su precio superaba los \$10.000 (Figura 5).

Figura 5. CamCube 3.0



- OCCIPITAL STRUCTURE SENSOR [8]

Structure Sensor, según describe su fabricante Occipital, es el primer sensor 3D para dispositivos móviles. Se acopla al dispositivo móvil (de momento dispositivos IOS) permitiendo capturar mapas 3D de espacios interiores y exteriores y almacenar cada medición en el móvil. Es posible capturar instantáneamente modelos 3D de objetos y personas que puedan ser importados a programas de diseño asistido por computador (CAD) para su impresión 3D. También es posible utilizarlo en aplicaciones de realidad aumentada en las que el mundo real es el mundo que te muestra la aplicación. Está optimizado en un rango de entre 40 cm a más de 3,5 metros. Occipital proporciona una librería conocida como Structure SDK que permite desarrollar aplicaciones para dispositivos IOS. Hasta el momento es la única vía para desarrollar aplicaciones, pero hay que tener en cuenta que el creador de la librería de código abierto OpenCV está implicado en el proyecto por ser consejero de la empresa Occipital. Su coste está en torno a los \$350 (Figura 6).

Figura 6. Occipital Structure Sensor



- SOFTKINETIC DEPTHSENSE [9]

SoftKinetic ha colaborado en diferentes proyectos basados en reconocimiento de gestos para diferentes plataformas como PC, consolas, arcade y señalización digital interactiva. Proporcionan diferentes dispositivos que permiten capturar mapas de profundidad para ser utilizadas en la construcción de aplicaciones para diferentes propósitos como uso en espacios interiores, fines automovilísticos, sistemas de atención médica inteligentes, etc. que realicen un seguimiento preciso de los dedos de la mano, de la propia mano e incluso de la totalidad del cuerpo. Disponen también de micrófonos que permiten capturar sonido y también acelerómetros que permiten medir posibles movimientos del dispositivo. Su rango de distancias de uso es de 15 cm a 4,5 metros. Son compatibles con librerías tales como IISU SDK de SoftKinetic, Intel Perceptual Computing SDK y OpenCV [10]. El precio de los diferentes dispositivos SoftKinetic DepthSense no superan los \$300 (Figuras 7 y 8).

Figura 7. SoftKinetic DepthSense 325



Figura 8. SoftKinetic DepthSense 311



- CREATIVE SENZ3D [11]

Este dispositivo permite detectar gestos de la mano y el movimiento de la cabeza, abriendo nuevas posibilidades de interacción. Captura imágenes de alta resolución 720p e incorpora dos micrófonos con reducción de ruido. Es posible realizar reconocimiento facial y reconocimiento de habla. Es posible desarrollar aplicaciones utilizando librerías como Intel Perceptual Computing SDK, IISU SDK de SoftKinetic, OpenCV y PCL. El rango de este dispositivo es de 15 cm a 1 metro. El precio del dispositivo es de \$200, lo que lo hace muy interesante por las características del dispositivo. Su principal inconveniente es su reducido rango de acción (Figura 9).

Figura 9. Creative Senz3D



- PANASONIC D-IMAGER [12]

Se trata de un sensor basado en la tecnología TOF que permite detectar gestos humanos y realizar un seguimiento completo del movimiento del cuerpo permitiendo una completa experiencia de interacción. Permite desarrollar aplicaciones de señalización digital, medición de audiencia, seguridad, atención médica, detección de formas y detección de personas en los accesos a recintos. Su rango de actuación es de 1,2 metros a 9 metros en función del modelo y permite desarrollar aplicaciones utilizando librerías como IISU SDK de SoftKinetic, Omek Beckon SDK, OpenCV, PCL y Flash. Uno de sus inconvenientes es el elevado coste, según el modelo, en torno a \$2.000 (Figura 10).

Figura 10. Panasonic D-IMager



- PRIMESENSE [13]

PrimeSense era una empresa de Israel especializada en detección 3D que ha sido comprada recientemente por Apple (Noviembre 2013) cuyo futuro es una incógnita ya que no han informado todavía hacia donde van a orientar su estrategia corporativa con respecto a los dispositivos de captura de mapas de profundidad. Sus dispositivos utilizaban tecnología infrarroja para producir mapas de profundidad. Permitía capturar video, profundidad y sonido y su diseño ha sido la base para que otras empresas como Asus y Microsoft crearan sus propios dispositivos. Se podían adquirir por un precio en torno a \$400 y su rango de actuación era de 80 cm a 3,5 metros. Permitía desarrollar aplicaciones utilizando librerías como OpenNI y NITE proporcionadas por PrimeSense, OpenCV y PCL (Figura 11).

Figura 11. PrimeSense Carmine



- ASUS XTION PRO [14]

Tal y como describe en su página web, ASUS Xtion PRO es el primer dispositivo de captura de movimiento exclusivo para ordenador. Dispone de sensores de color, profundidad y sonido que permite desarrollar aplicaciones de distinto ámbito utilizando las librerías OpenCV, OpenNI y NITE. Su rango de actuación es de 80 cm a 3,5 metros y su precio es en torno a 150€. Con la compra de PrimeSense y su librería de desarrollo NITE por parte de Apple es una incógnita la disponibilidad de dichas librerías y sus posibilidades de desarrollo (Figura 12).

Figura 12. ASUS Xtion PRO Live



- SONY PLAYSTATION CAMERA [15]

Sony proporciona dispositivos propios de captura de movimiento dirigidos a su uso en las consolas PlayStation. Dichos dispositivos han ido evolucionando hasta obtener mapas de profundidad que podrían ser utilizadas para el desarrollo de aplicaciones cuyo objetivo es el seguimiento del movimiento del cuerpo humano. Sin embargo sus características son similares a las de sensores de otros fabricantes y no han publicado librerías potentes que permitan el desarrollo de aplicaciones en sus dispositivos. El precio de su dispositivo más moderno, Sony PlayStation Camera es de \$60. Es posible utilizar la librería CL Eye Platform SDK, ya obsoleta, IISU SDK y OpenCV. Uno de sus inconvenientes es el escaso soporte que proporciona Sony a los desarrolladores y, por tanto, el poco uso de sus dispositivos en la creación de nuevas aplicaciones (Figuras 13 y 14).



- MICROSOFT KINECT [16]

Kinect es un dispositivo de captura de movimiento que puede ser utilizado tanto en las consolas Xbox de Microsoft como en aplicaciones de ordenador. Está basado en el diseño de PrimeSense que utiliza tecnología infrarroja y permite el reconocimiento de gestos y habla, controlando aplicaciones con el cuerpo y la voz, además de realizar el seguimiento del cuerpo en 3 dimensiones. Su rango de actuación es de 80 cm a 4 metros. Su precio en su versión de ordenador es en torno a \$250 aunque es posible utilizar su versión de consola que vale en torno a 100\$. Por otra parte Microsoft está a punto de sacar al mercado una nueva versión de Kinect (2.0) que mejorará la resolución de las imágenes de color, profundidad y también el

audio. Es posible desarrollar aplicaciones con la librería que ofrece Microsoft, además de poder utilizar librerías como OpenNI, OpenCV, IISU SDK, etc. Se puede afirmar que Microsoft facilita el desarrollo de aplicaciones ya que su dispositivo tiene compatibilidad con la mayoría de librerías más potentes, además de proporcionar una librería propia con una importante cantidad de funcionalidades que permite programar en lenguajes como C++, C# y Visual Basic. Es por esto que hay un gran número de proyectos que utilizan Kinect como base para realizar aplicaciones en una amplia variedad de ámbitos (Figuras 15 y 16).

Figura 15. Microsoft Kinect



Figura 16. Microsoft Kinect 2.0



2.2.4. Características y tipología del instrumento seleccionado.

Una vez estudiadas en profundidad las características y las posibilidades que brindan cada uno de los instrumentos mencionados, los cuales nos permiten generar mapas de profundidad para registrar el movimiento del alumnado, hemos escogido el dispositivo Kinect de Microsoft debido, fundamentalmente, a los siguientes factores:

- Su precio hace de él un producto accesible.
- Disponibilidad de algunos de estos dispositivos en la UA para préstamo.
- Genera mapas de profundidad con buena resolución para su análisis.
- Compatibilidad con diversas librerías, tanto de pago como de uso gratuito, que permiten analizar el movimiento de sujetos.
- Soporte por parte de Microsoft con una sólida librería en continua mejora y de una amplia comunidad de desarrolladores.

3. CONCLUSIONES

Tras hacer una búsqueda del trabajo realizado previamente y la literatura existente en este ámbito observamos que es un campo de trabajo con enormes posibilidades y que, al mismo tiempo, está poco desarrollado.

Hoy en día disponemos de múltiples dispositivos que utilizan la tecnología de generación de mapas de profundidad, que permiten la captura de movimiento de sujetos.

Además, estos instrumentos cumplen los objetivos iniciales de este trabajo pues, por una parte son capaces de capturar datos manteniendo el anonimato de los sujetos estudiados y, por otra, se pueden introducir en el contexto de investigación (aula), aumentando enormemente la validez ecológica de los datos recogidos.

4. REFERENCIAS

4.1 Referencias bibliográficas.

- García, R., Pérez, F., Pérez, & Natividad, L. (2012). Evaluación del estrés académico en estudiantes de nueva incorporación a la universidad. *Revista latinoamericana de psicología*, vol. 44 (2), pp. 143-154.
- León, B. (2008). Atención plena y rendimiento académico en estudiantes de enseñanza secundaria. *European journal of education and psychology*, vol. 1 (3), pp. 17-26.
- Requena, F. (1998). Género, redes de amistad y rendimiento académico. *Revista de sociología*, vol.56, pp. 233-242.
- Thalmann, D., Gutiérrez, M. & Vexo, F. (2008). *Steppin intro Virtual Reality*. Springer-Verlag.
- Hoffman, M., Granter, N., Witte, K., Edelman-Nusser, J. & Nowoisky, C. (2006). *Use of the infrared based motion capture system as 200 in sport science*. En *The Engineering of Sport 6*, pp. 45-50. Springer, New York.
- Kanetaka, H., Yabukami, S., Hashi, S. & Arai, K. (2010). *Wireless magnetic motion capture system for medical use*. En Sansano, T. y Suzuki, O., editores, *Interface Oral Health Science 2009*, pp. 329-331. Springer, Japan.
- Leow, W., Wang, R. & Leong, H (2012) *3-d-2-d spatiotemporal registration for sports motion analysis*. *Machine Vision an Applications*, vol. 23 (6), pp. 1177-1194
- Sanmartín, G., Flores, J., Arias, P., Cudeiro, J. & Méndez, R. (2012). *Motion capture for clinical purposes, an approach using primesense sensors*. En Perales, F., Fisher, R., y Moeslund, T. (ed.). *Articulated Motion and Deformable Objects*, vol. 7378 de *Lecture Notes in Computer Science*, pp. 273-281. Springer, Berlín: Heidelberg.

4.2 Referencias web.

- [1] Ley Orgánica 1/1982, de 5 de mayo, sobre protección civil del derecho al honor, a la intimidad personal y familiar y a la propia imagen. http://noticias.juridicas.com/base_datos/Admin/lo1-1982.html.

- [Consulta: 10 de mayo de 2014].
- [2] Tracker Video Analysis and Modeling Tool.
<http://www.opensourcephysics.org/items/detail.cfm?ID=7365>.
[Consulta: 18 de mayo de 2014].
- [3] Kynovea. <http://kynovea.org>. [Consulta: 18 de mayo de 2014].
- [4] Darthfish. <http://www.dartfish.com/en/>. [Consulta: 23 de mayo de 2014].
- [5] Laboratorio de Creación de Contenidos Multimedia.
<http://biblioteca.ua.es/es/fragua/recursos/laboratorio-de-creacion-de-contenidos-multimedia.html>
[Consulta: 27 de mayo de 2014].
- [6] SwissRanger TOF Cameras. <http://www.mesa-imaging.ch/products/product-overview/>
<http://en.souvr.com/product/201005/6505.html> [Consulta: 27 de mayo de 2014].
- [7] PMD CamCube. http://www.pmdtec.com/news_media/video/camcube.php [Consulta: 27 de mayo de 2014].
- [8] Occipital Structure Sensor. <http://occipital.com/> [Consulta: 27 de mayo de 2014].
- [9] SoftKinetic DepthSense Cameras. <http://www.softkinetic.com/en-us/products/depthsensecameras.aspx> [Consulta: 28 de mayo de 2014].
- [10] Middlewares (SDKs) y dispositivos.
<http://www.softkinetic.com/products/iisummiddleware.aspx> [Consulta: 28 de mayo de 2014].
- [11] Creative Senz3D Gesture Camera. <http://us.creative.com/p/web-cameras/creative-senz3d>
[Consulta: 28 de mayo de 2014].
- [12] Panasonic D-IMager. <http://pewa.panasonic.com/components/built-in-sensors/3d-image-sensors/d-imager/> [Consulta: 28 de mayo de 2014].
- [13] PrimeSense. <http://en.wikipedia.org/wiki/PrimeSense> [Consulta: 28 de mayo de 2014].
- [14] ASUS Xtion Pro. <http://www.asus.com/News/RsnFvznTgaEXpJzR> [Consulta: 29 de mayo de 2014].
- [15] Playstation4 Camera.
<http://us.playstation.com/ps4/features/techspecs/index.htm?sel=techspecs-camera>
[Consulta: 29 de mayo de 2014].
- [16] Kinect for Microsoft. <http://www.microsoft.com/en-us/kinectforwindows/> [Consulta: 29 de mayo de 2014].