

Desarrollo de un Entrenador Virtual de Capoeira

Víctor Tobes¹, Raúl Fernández¹, Gonzalo Méndez^{1,2}, and Pablo Gervás^{1,2}

¹ Facultad de Informática, Universidad Complutense de Madrid

² Instituto de Tecnología del Conocimiento, Universidad Complutense de Madrid
`{raulfe03,vtobes,gmendez,pgervas}@ucm.es`

Resumen La investigación en *Reactive Virtual Trainers (RVT)* basados en el uso de MS Kinect ha dado lugar a una diversidad de soluciones para la danza, las artes marciales y el ejercicio físico en general. Este artículo presenta un prototipo de sistema para el entrenamiento de capoeira, un arte marcial brasileño a caballo entre los deportes de lucha y la danza. Este entorno está pensado para ser utilizado por el alumno para que pueda aprender a realizar diferentes movimientos de capoeira sin la necesidad de que esté presente un profesor. El entrenamiento consiste en imitar una serie de movimientos, los cuales han sido previamente capturados por varios expertos en el arte marcial. El entrenamiento virtual se desarrolla acorde al nivel que posea el alumno, ya sea principiante, aprendiz o avanzado. De este modo, el alumno va aprendiendo a realizar los movimientos de forma progresiva.

1. Introducción

La aparición de hardware relativamente asequible para la captura de movimiento en los últimos años ha dado lugar a un crecimiento del número de soluciones para el entrenamiento virtual de ejercicios físicos. Estas soluciones combinan la captura de movimiento con la representación en entornos 3D para permitir al usuario realizar movimientos físicos, verlos representados en pantalla, y recibir realimentación sobre ellos. Estos sistemas se denominan *Reactive Virtual Trainers (RVT)*.

La capoeira es un arte marcial afro-brasileño que combina facetas de danza, música y acrobacias, así como expresión corporal [1]. Fue desarrollado en Brasil por descendientes africanos con influencias indígenas, probablemente a principios del siglo XVI. Es conocido por sus rápidos y complejos movimientos, que utilizan distintas partes del cuerpo (brazos, piernas, codos, rodillas, cabeza...) para ejecutar maniobras de gran agilidad en forma de patadas, fintas y derribos, entre otros. La capoeira como estilo de lucha incorpora movimientos bajos y barridos, mientras que en el ámbito deportivo se hace más énfasis en las acrobacias y las demostraciones ritualizadas de habilidad. Se practica con música tradicional de berimbau.

Una de las características fundamentales de la capoeira, como de muchas otras artes marciales, consiste en la realización de manera repetida de un mismo movimiento para ir mejorando paulatinamente su ejecución, tanto en su trayectoria como en su velocidad o su fuerza. Sin embargo, especialmente en el caso

de los principiantes, resulta complejo apreciar si existe mejora sin la supervisión constante de un entrenador, lo cual no resulta posible en la mayor parte de las ocasiones. Por este motivo, se ha planteado el desarrollo de un entrenador de capoeira, inicialmente en un nivel básico, que, mediante la utilización de tecnologías que ya son de uso doméstico (como son un ordenador personal y **Kinect**), supervise el proceso de entrenamiento de una persona que se inicie en el aprendizaje de la capoeira.

2. Trabajo Relacionado

Para la realización de este trabajo se han revisado trabajos anteriores en el área de los *Reactive Virtual Trainers* y representación de personajes animados en entornos 3D.

2.1. Reactive Virtual Trainers

Los *Reactive Virtual Trainers (RVT)* es una tecnología que se viene utilizando desde hace algunos años para la danza, las artes marciales y los ejercicios físicos en general y la rehabilitación en particular.

Los sistemas de representación digital de danza existentes se han centrado esencialmente en la representación realista de la danza, y hay poco trabajo sobre maneras de presentar realimentación al alumno cuando son utilizados para enseñanza [5].

El sistema [5] utiliza **Kinect** para registrar los movimientos de un bailarín de ballet y representa el movimiento como un *espacio de posturas* en forma de trayectorias de gestos. Un sistema de reconocimiento permite comparar el movimiento del alumno con los registrados para el profesor.

Con respecto a las artes marciales el trabajo de [4] es un maestro virtual de Kung Fu que ayuda a los usuarios a aprender y perfeccionar distintas técnicas basándose en la comparación entre el movimiento realizado por el usuario y movimientos pregrabados realizados por el maestro. Este sistema utiliza para comparar las dos versiones del movimiento un algoritmo de *Dinamic Time Warping* basado en la distancia euclídea.

Para el caso de ejercicios de educación física y rehabilitación se ha trabajado más en el desarrollo de métodos de realimentación. En [14] se plantean soluciones específicas basadas en aprendizaje máquina para clasificar los movimientos del alumno (en el contexto de ejercicios de educación física) en correctos o incorrectos. The OnlineGym [12] presenta un mundo virtual dirigido a usuarios de edad avanzada con necesidades específicas de ejercicio, en el que los usuarios pueden participar mediante el uso de un sistema de captura de movimiento. El trabajo de [13] modela de manera más elaborada a un entrenador, proporcionando no sólo realimentación sobre la corrección de los ejercicios sino también proponiendo secuencias adecuadas de ejercicios y ayudando a motivar al alumno.

2.2. Representación y Animación 3D

El software principal utilizado en este proyecto es **Unity 5.5** [16], ya que se trata de un motor para el desarrollo de videojuegos muy versátil. Con el fin de elaborar, compilar y depurar el proyecto se ha utilizado **Visual Studio 2015** [8] y **MonoDevelop** [11] ya que mantienen una sola base de código para todas las plataformas. Respecto a la creación de los personajes, previamente se utilizó la aplicación **MakeHuman** [6], pero se observó que **Fuse Character Creator** de **Mixamo** [9] ofrecía una creación de modelos humanos más estéticos. En cuanto a las animaciones, se hizo uso de **Mixamo 3D Animation Online Services** [10], ya que es una compañía tecnológica encargada del desarrollo y venta de servicios para la animación de personajes 3D. A través de **Marvelous Designer** [7] se probó la confección de la ropa que utilizarían los avatares, pero se optó por utilizar los diseños predefinidos de **Fuse Character Creator** de **Mixamo** debido a la alta calidad de las texturas. **Blender** [3] se utilizó para modelar y renderizar el logo de la escuela **Abadá-Capoeira** en las camisetas de los personajes.

3. Un Entrenador Virtual de Capoeira

El entrenador virtual de capoeira desarrollado puede utilizarse en modo profesor (para grabar las secuencias de referencia de cada ejercicio) o en modo alumno (para entrenar secuencias concretas asesorado por el sistema). En este apartado se presenta la funcionalidad básica del entrenador de capoeira en ambos modos, el mecanismo de comparación de los movimientos realizados por el alumno frente a los del profesor y el proceso de grabación de movimientos. Para una descripción más extensa del sistema puede consultarse la información proporcionada en [15].

3.1. Funcionalidad Básica del Alumno

Si el usuario es un alumno, se entrará en la escena principal del proyecto, donde se desarrolla toda la funcionalidad del entrenamiento virtual. En el interior de la escena se muestra la misma interfaz que en el caso del profesor, con la diferencia de que este no podrá grabar movimientos para su posterior análisis.

Considerando que al comenzar a formarse en un arte marcial se desconocen los nombres de los movimientos que se desean poner en práctica, se le ofrece al alumno la posibilidad de visualizar previamente los movimientos antes de empezar la práctica para poder seleccionar más fácilmente el que desea. Para iniciar el entrenamiento acorde a su experiencia, el alumno tendrá que seleccionar el nivel que le corresponda, ya sea principiante, aprendiz o avanzado.

Para iniciar la práctica de un movimiento, el alumno deberá adoptar una pose inicial correspondiente al movimiento seleccionado previamente. Hasta que el alumno no tome esa misma pose, el sistema estará esperando la calibración. En el instante en que la pose del alumno sea la misma que la del entrenador virtual, el sistema estará calibrado y empezará una cuenta atrás de tres segundos, tras la cual se empezará a reproducir el movimiento seleccionado.



Figura 1: Animación de movimiento realizado

En el momento en que el entrenador virtual inicie una reproducción de un movimiento, el alumno deberá imitarlo lo más fielmente posible. Mientras se realizan las transiciones, el sistema comparará las posición de los 25 *joints* que componen el esqueleto del entrenador virtual, con las posiciones de los 25 *joints* correspondientes al alumno capturado por Kinect.

Tras finalizar el movimiento, el asistente mostrará una animación de éxito en el caso de que el movimiento se haya realizado de forma correcta (ver Figura 1), o una animación de fracaso en el caso de que se produzca algún error en la ejecución del movimiento. Posteriormente, se presenta un esqueleto en el que se muestran de color rojo las partes del cuerpo cuya colocación debe ser corregida, y en color verde las que se han utilizado de manera correcta (ver Figura 6). A demás, se visualizará una serie de indicaciones en forma de texto para que el alumno pueda corregir la posición errónea de una forma mas efectiva (ej. *Gira el brazo izquierdo hacia atrás*).

3.2. Funcionalidad Básica del Profesor

Cuando el usuario inicia el sistema como profesor, puede acceder a la funcionalidad de grabación de nuevos movimientos. Al comenzar a realizar la captura es necesario introducir el nombre deseado para el nuevo movimiento y colocarse a unos tres metros de Kinect para ser capturado correctamente. A partir de ese momento, el sistema graba los movimientos realizados por el usuario y los guarda en ficheros de texto para poder utilizarlos tanto para mostrárselos a los alumnos como para compararlos con los movimientos de estos cuando se encuentren practicando. Con el fin de obtener una mayor inmersión en la aplicación, se utiliza el modo espejo como efecto para reproducir los movimientos captados.

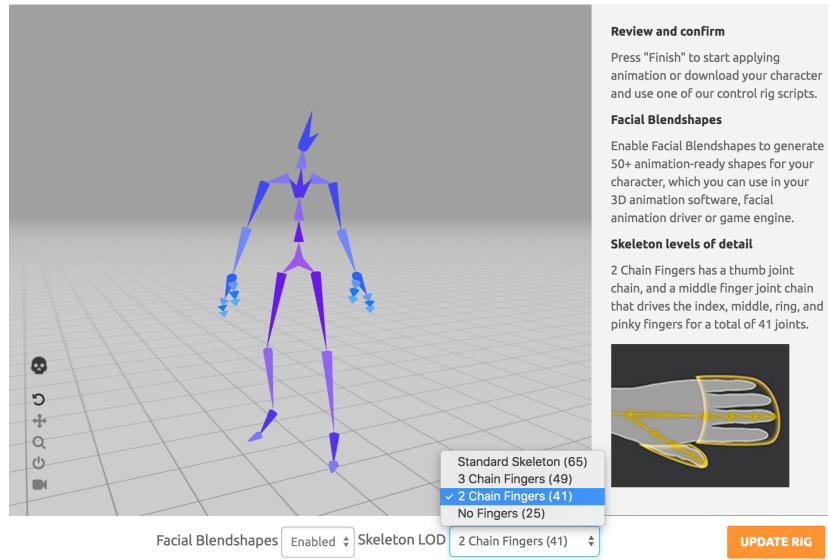


Figura 2: Representación utilizada para el avatar del usuario

El profesor también puede utilizar la funcionalidad del alumno, ya que será necesario revisar y reproducir el correcto funcionamiento de los nuevos movimientos capturados con Kinect. Inicialmente, la aplicación incluye ocho movimientos capturados a instructores de la escuela **Abadá-Capoeira**.³

3.3. Avatares, Animaciones y Escenarios

Se han desarrollado los diferentes avatares con una fisionomía humana para dar una sensación más agradable a los *cubeman*⁴ que representan a los usuarios y de este modo se proporciona una perspectiva más amena al mismo.

A fin de realizar un entorno más realista se crearon cuatro avatares diferentes, cada uno con una fisionomía propia. A la hora de elaborar el vestuario de los personajes, se creó una ropa muy similar a la utilizada por la escuela **Abadá-Capoeira**: se usaron prendas prediseñadas en la aplicación **Fuse Character Creator** como los pantalones, camisetas y el top que utilizarían los avatares del sistema con los retoques apropiados para darle ese color blanco característico.

Los dos avatares principales, como son el alumno y el profesor, se desarrollaron con la misma vestimenta que utilizaba la escuela **Abadá-Capoeira**. En cambio, los dos avatares que se emplean para dar un mayor dinamismo a la aplicación utilizan el mismo pantalón largo, mientras que el avatar masculino no utilizará prenda superior y el avatar femenino viste un top en color blanco.

³ <http://www.abadacapoeira.com.br/>

⁴ Esqueleto verde que emula el movimiento de la persona captada.

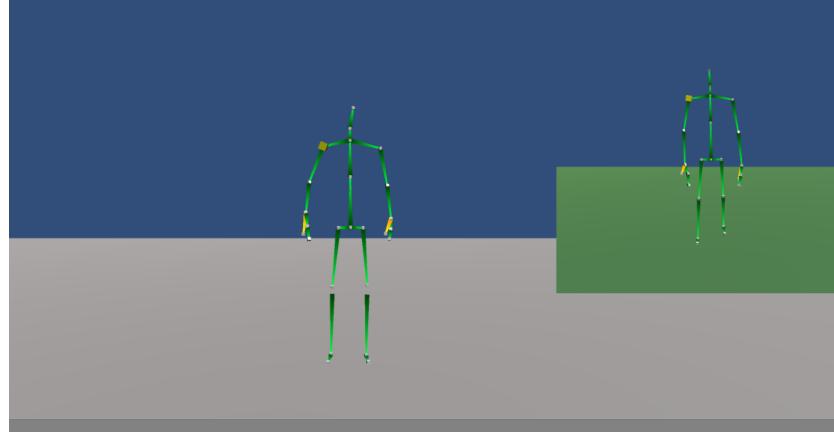


Figura 3: Ilustración de los *cubeman*

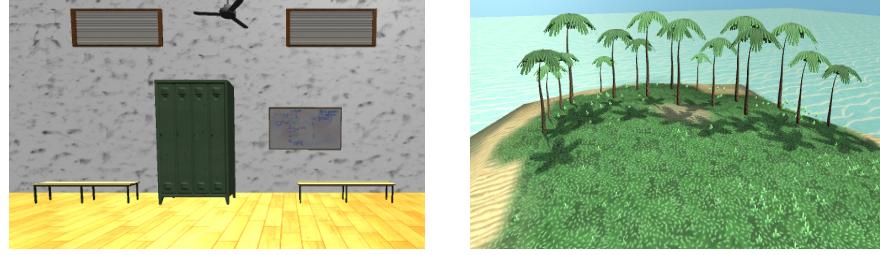
Tras completar la estética de los personajes, **Fuse Character Creator** de **Mixamo** ofrece la posibilidad de configurar los huesos de los avatares para posteriormente utilizar las animaciones que suministra **Mixamo**. El avatar elegido que mejor se adapta a **Kinect** es el que contiene 41 huesos, ya que demuestra una movilidad más semejante al *cubeman* inicial. El conjunto de huesos en cuestión aparece representado en la Figura 2.

La representación del movimiento del usuario se muestra mediante un *cubeman*, que simula todo el movimiento que realiza el usuario. Los datos del *cubeman* (ver Figura 3) están definidos con 25 **GameObjects** llamados **Joints**, que hacen una representación de las articulaciones del cuerpo humano. Estos **Joints** se representan en **Unity** con unas coordenadas (X,Y,Z) y una rotación. Toda la información del movimiento se obtiene a partir del sensor de **Kinect**.

La información generada en cada *frame* por cada *cubeman* es clasificada y guardada en un fichero de texto para su futura utilización, como el seguimiento del progreso del alumno a lo largo del tiempo. A su vez, se ha creado una funcionalidad para poder reproducir un movimiento grabado a la vez que muestra al usuario captado en tiempo real.

Para incorporar las animaciones al avatar, se crea el componente **Animator Controller** de **Unity**, encargado de añadir y gestionar animaciones. Este componente contiene una máquina de estados capaz de coordinar que animación se ejecuta en cada momento. El estado inicial será la *ginga*, y se realizará una transición a la animación de victoria o derrota que dependerá de la comparación del movimiento del usuario con el del profesor. Una vez determinado el estado, este volverá a la animación de *ginga* para continuar con el siguiente movimiento.

Para el desarrollo de los escenarios, se han creado dos ambientes diferentes: uno escenificando un entorno de entrenamiento como es el caso de un gimnasio para la visualización del alumno (Figura 4a) y otro más orientado al profesor y los animadores, como se trata de una playa Brasileña (Figura 4b).



(a) Alumno

(b) Profesor y animadores

Figura 4: Escenarios

3.4. Comparación de los Movimientos del Usuario y del Profesor

La comparación del movimiento se realiza mediante el reconocimiento de un gesto de *Kinect*. La funcionalidad del gesto se ha implementado utilizando estados y progresión, permaneciendo en el primer estado hasta que identifique una posición específica. Para completar el gesto se tendrá un tiempo determinado para llegar a la progresión final del movimiento. En el caso de que ese tiempo se termine, el gesto se cancelará haciendo que fracase la comparación.

El gesto fue nombrado como *Move* y está formado por dos estados (ver Figura 5). El estado inicial se encarga de calibrar la posición inicial y el siguiente estado se define como el estado por defecto, encargado de la comparación de la lista de los movimientos. La comparación consiste en igualar las coordenadas X e Y de los *Joints* con un margen de error modificable por el usuario, y si cumple este requisito, pasará al siguiente estado. Permanecer en el estado por defecto tiene un tiempo límite, que será el tiempo del movimiento más dos segundos de espera. En ese tiempo el usuario deberá superar todos los estados para que el movimiento se considere como realizado correctamente.

A su vez, para realizar el análisis se guarda toda la información del movimiento en el estado por defecto del gesto, hasta donde el usuario falle el movimiento. Cuando la comparación termina se podrá activar el análisis, haciendo que los avatares se desplacen a primera escena y modificándolos para que se vean transparentes. De este modo se creará un esqueleto verde parecido a los *cubeman* que representan a los usuarios. Los avatares simularán el movimiento hasta el *frame* donde se equivocó el usuario y se realizará una comprobación de cada *Joint* para cambiar el hueso de color de verde a rojo (Figura 6) para mostrar los fallos.

A la vez, se va generando una cadena de texto con una corrección de posicionamiento compuesto por tres partes. La primera es la parte del cuerpo que estamos corrigiendo. La segunda parte está formada por la elección de la dirección, la cual se comprueba realizando la diferencia de la coordenadas X del *Joint* del usuario, con la del *Joint* de la grabación. Si la diferencia es positiva será hacia la derecha y en el caso de ser negativa será hacia la izquierda. Para la tercera parte se realiza la diferencia con la coordenada Y, definiendo así la

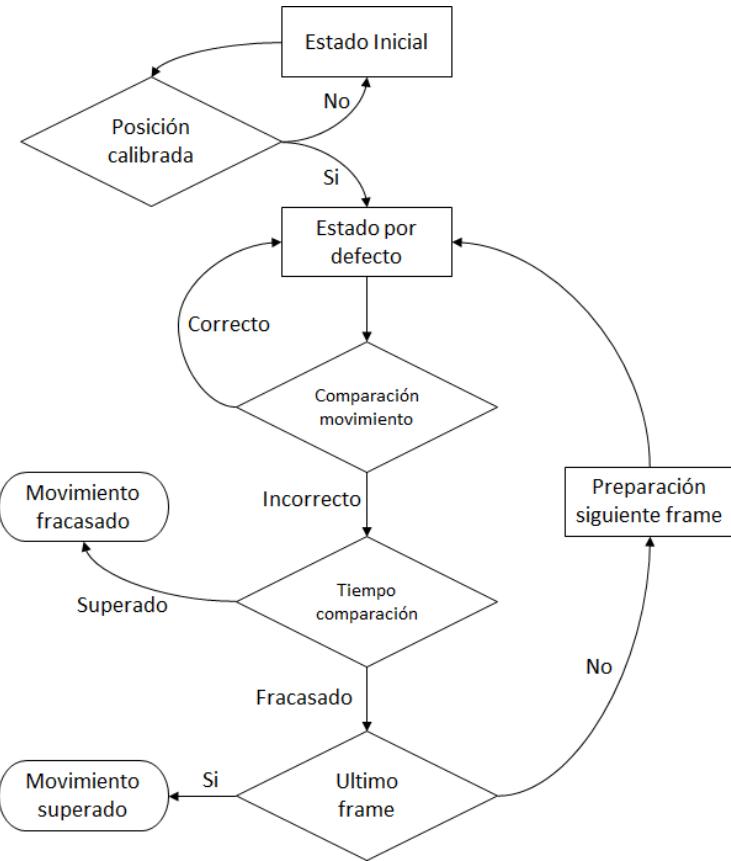


Figura 5: Diagrama del gesto *Move*

altura. Si la diferencia es positiva la altura sera hacia abajo, en caso contrario sera hacia arriba.

3.5. Grabación de Movimientos con Expertos

En esta etapa del proyecto se visitó la Asociación Cultural Deportiva Rio⁵, donde se reúnen los practicantes expertos en capoeira, para realizar una serie de grabaciones y pruebas. Este proceso consistía en grabar una batería de movimientos de capoeira para seleccionar los captados correctamente por Kinect. Para facilitar la grabación al profesor de capoeira, se optó por capturar varias repeticiones del mismo movimiento en la misma grabación. De este modo, y con la supervisión del profesor, se tuvieron que seleccionar los movimientos que mejor se asemejaban al real, recortando el intervalo de *frames* más adecuado para incluirlo en la aplicación.

⁵ <https://www.asociacionrio.com/>

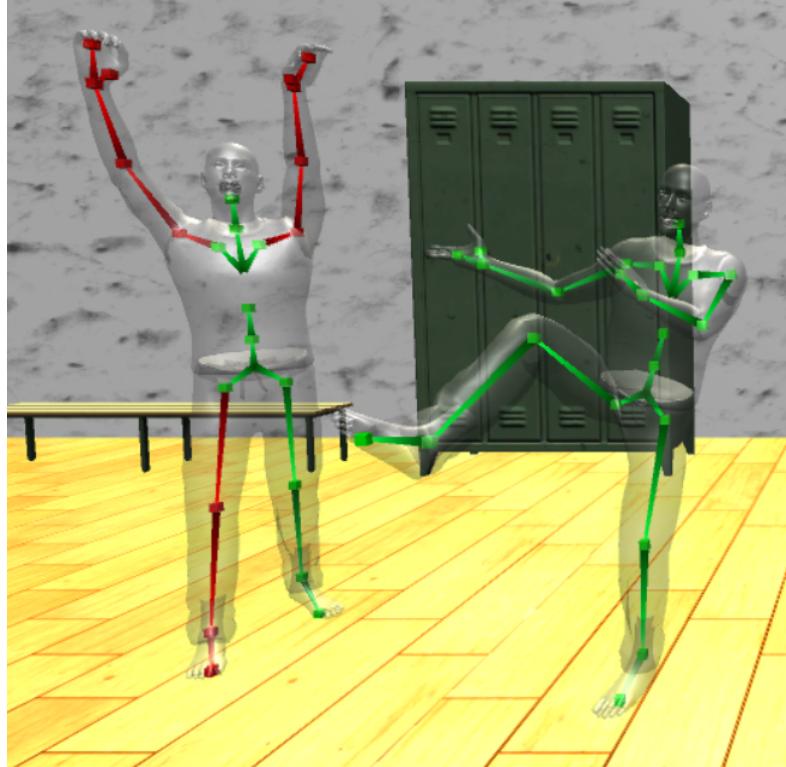


Figura 6: *Feedback* del análisis

De los 20 movimientos grabados, 8 fueron factibles para su uso, mientras que el resto, al incluir giros, presentaron problemas en la grabación. Esto se debe a que *Kinect* sólo posee una cámara frontal, de modo que al rotar alrededor del eje vertical se acaba perdiendo la referencia de cuáles son las extremidades izquierda y derecha. Como resultado, al darse la vuelta, la captura vuelve a mostrar al usuario de frente, pero con los brazos y las piernas cruzadas.

4. Discusión

Como se mencionó en la introducción, uno de los objetivos al desarrollar este sistema era que pudiese funcionar con tecnología de uso doméstico, lo cual restringe notablemente el abanico de posibilidades a considerar. La configuración elegida para el sistema utiliza una única cámara para la captación del movimiento del usuario, proporcionada por *Kinect*. Esto limita mucho la capacidad del sistema de percibir los movimientos en toda su complejidad, sobre todo en aquellos casos en que hay rotación del torso, lo cual puede generar confusiones entre distintos miembros del avatar que tengan simetría axial (tras el giro el sistema



Figura 7: Aspecto final de la aplicación

confunde un brazo con otro, y piensa que el usuario sigue de frente cuando en realidad está de espaldas). Lo mismo sucede a la hora de analizar la profundidad de algunos de los movimientos realizados. Se espera en el futuro poder ampliar este trabajo con la utilización de mayor número de cámaras.

Por otro lado, algunos de los trabajos estudiados (e.g. [4,5]) utilizan técnicas más sofisticadas para analizar los movimientos de los usuarios que las usadas en este trabajo. Si bien esto puede ser útil en niveles de aprendizaje avanzados, con principiantes en la práctica de la capoeira las diferencias con los movimientos realizados por el profesor son poco sutiles, con lo que un método más sencillo de comparación de movimientos resulta, por un lado, menos complejo y más rápido de calcular, y por otro, el resultado de la comparación resulta más fácilmente traducible a texto para proporcionar al aprendiz explicaciones útiles acerca de cómo mejorar la realización de las técnicas.

El trabajo realizado se enmarca en el contexto del proyecto IDiLyCo: Inclusión Digital, Lenguaje y Comunicación (TIN2015-66655-R), en el que se espera poder dar aplicación a las tecnologías básicas desarrolladas para construir herramientas para la asistencia a personas con necesidades especiales.

5. Conclusiones y Trabajo Futuro

A lo largo del presente artículo se ha descrito el desarrollo de un entrenador virtual de capoeira basado en la captura de los movimientos del usuario a través de Kinect y en su comparación con los movimientos grabados previamente de un profesor de capoeira, con la consiguiente explicación de cómo conseguir realizar correctamente los movimientos ejecutados de manera errónea.

Aunque aún no se ha realizado una evaluación extensa con practicantes novatos y expertos de capoeira, sí se han realizado pruebas preliminares para comprobar que la aplicación funciona correctamente y que los comentarios proporcionados para corregir los errores son adecuados. Estas pruebas se han realizado con

una muestra pequeña de usuarios de distintos niveles y los resultados preliminares muestran que, al nivel al que se han realizado las pruebas, y con movimientos no excesivamente complejos de cuerpo entero, el funcionamiento de la aplicación es rápido y los resultados de los análisis y las explicaciones de los errores se acercan bastante a la realidad, por lo que actualmente nos hallamos preparando una evaluación más extensa de la aplicación.

Sin embargo, como se ha mencionado en la discusión, la utilización de una única cámara para captar los movimientos no permite capturar correctamente movimientos que impliquen giros sin utilizar elementos adicionales, como marcadores. Por lo tanto, nos planteamos una solución similar a la descrita en [2] donde utilizan dos Kinect para solucionar este problema, aunque esta opción aleja la solución de su uso doméstico.

Por otro lado, se plantea también la necesidad de incluir métodos más sofisticados para el análisis de los movimientos, similares a los usados en [4,5] para permitir un análisis más fino de los movimientos que pueda proporcionar resultados de mayor utilidad para practicantes avanzados de capoeira.

Finalmente, también se contempla el análisis de los movimientos de más de un usuario de manera simultánea, de forma que se puedan reproducir situaciones similares al trabajo por parejas que tiene lugar en una *roda* de capoeira, donde se ponen en práctica técnicas de combate cuerpo a cuerpo.

Agradecimientos

Este proyecto ha sido financiado por el proyecto IDiLyCo: Inclusión Digital, Lenguaje y Comunicación, Ref. No. TIN2015-66655-R (MINECO/FEDER). Los autores quieren agradecer al Instructor Gato y al Graduado Windows, de la escuela Abadá Capoeira⁶ y la Asociación Cultural Deportiva RIO⁷ de Madrid por su colaboración y asesoría en la grabación de las versiones de referencia de los ejercicios.

Referencias

1. Assunção, M.: Capoeira: A History of an Afro-Brazilian Martial Art. Routledge (2005)
2. Gao, Z., Yu, Y., Zhou, Y., Du, S.: Leveraging two kinect sensors for accurate full-body motion capture. Sensors 15(9), 24297–24317 (2015)
3. <https://www.blender.org/>: Blender. [online] [Accessed 31 May 2017].
4. Keerthy, N.K.: Virtual Kung fu Sifu with Kinect. Master's thesis, San José State University, CA, USA (2012)
5. Kyan, M., Sun, G., Li, H., Zhong, L., Muneesawang, P., Dong, N., Elder, B., Guan, L.: An Approach to Ballet Dance Training Through MS Kinect and Visualization in a CAVE Virtual Reality Environment. ACM Trans. Intell. Syst. Technol. 6(2), 23:1–23:37 (Mar 2015)

⁶ <http://www.abadacapoeira.com.br/>

⁷ <https://www.asociacionrio.com/>

6. MakeHuman: MakeHuman - Open source tool for making 3D characters. [online] . [Accessed 31 May 2017]. <http://www.makehuman.org/>
7. MarvelousDesigner: Marvelous Designer. [online] . [Accessed 31 May 2017]. <https://www.marvelousdesigner.com/>
8. Microsoft: Microsoft Visual Studio. [online] . [Accessed 31 May 2017]. <https://www.visualstudio.com/>
9. Mixamo: Adobe Fuse CC - Make customized 3D models to power your designs. [online] . [Accessed 31 May 2017]. <https://www.mixamo.com/fuse>
10. Mixamo: Mixamo 3D Animation Online Services - Animated 3D characters. No 3D knowledge required.. [online] . [Accessed 31 May 2017]. <https://www.mixamo.com/>
11. Monodevelop: Monodevelop - Cross platform IDE for C#, F# and more. [online] . [Accessed 31 May 2017]. <http://www.monodevelop.com/>
12. Paredes, H., Cassola, F., Morgado, L., de Carvalho, F., Ala, S., Cardoso, F., Fonseca, B., Martins, P.: Exploring the Usage of 3D Virtual Worlds and Kinect Interaction in Exergames with Elderly, pp. 297–300. Springer International Publishing, Cham (2014)
13. Ruttakay, Z., van Welbergen, H.: Elbows Higher! Performing, Observing and Correcting Exercises by a Virtual Trainer, pp. 409–416. Springer Berlin Heidelberg, Berlin, Heidelberg (2008)
14. Staab, R.: Recognizing specific errors in human physical exercise performance with Microsoft Kinect. Master's thesis, California Polytechnic State University, CA, USA (2014)
15. Tobes, V., Fernández, R.: Entramiento de actividades físicas mediante Kinect. Trabajo de fin de grado, Facultad de Informática, Universidad Complutense de Madrid (2017)
16. Unity-Technologies: Unity - Manual: Unity Manual. [online] . [Accessed 07 June 2017]. <http://docs.unity3d.com/Manual/index.html> (2017)