

Espejo Aumentado: sistema interactivo de Realidad Aumentada basado en Kinect.

Lucía Vera, Jesús Gimeno, Inmaculada Coma, Marcos Fernández
Instituto de Robótica. Universitat de València.
{Lucia.Vera, Jesus.Gimeno, Inmaculada.Coma, Marcos.Fernandez}@uv.es

Resumen. Presentamos en este artículo un sistema de interacción basado en Realidad Aumentada que permite conversar a un personaje virtual controlado por un actor en tiempo real con el público asistente a través de una pantalla de gran formato a modo de espejo aumentado. El sistema integra imágenes de vídeo reales con las imágenes virtuales del avatar y otros objetos animados incorporados como apoyo a la conversación. Para su implementación hemos optado por la exploración y combinación de tecnologías: dos sistemas kinect para captura de movimientos corporales, mapa de profundidad e imágenes reales, un giróscopo para detección de movimiento de la cabeza y algoritmos de control para gestionar las expresiones y emociones del avatar.

Keywords: Realidad Aumentada, Captura de movimiento, Personajes Virtuales

1 Introducción y trabajo previo.

En el ámbito de la interacción persona-ordenador, la posibilidad de combinar el mundo real con información virtual, la capacidad de incorporar la participación de personajes virtuales en escenarios reales y la interacción del público tanto con dichos personajes como con otros objetos virtuales de la escena, son elementos que cada vez permiten mayor riqueza y complejidad en simulaciones, presentaciones e incluso actos públicos.

En este entorno es donde enmarcamos nuestro sistema de interacción basado en Realidad Aumentada (R.A.)¹, que permite conversar a un personaje virtual controlado por un actor en tiempo real con el público asistente a través de una pantalla de gran formato a modo de *espejo aumentado*. El sistema integra imágenes de vídeo reales con las imágenes virtuales del avatar y otros objetos animados incorporados como apoyo a la conversación. Hemos optado por un sistema basado en R.A. frente a otro tipo de presentaciones que utilizan animación o sistemas de Realidad Virtual por el mayor grado de integración entre el público y el avatar, que aparece dentro del espejo aumentado junto al personaje virtual.

¹ Este proyecto ha sido promovido y financiado por la Diputación de Valencia, para su utilización en la Feria de Turismo de Madrid (Fitur2011). En su desarrollo ha colaborado la empresa Creta, el grupo Artec del Instituto de Robótica para el desarrollo tecnológico y la empresa Hampa Animation Studio para el diseño gráfico.

Nuestra aplicación ha tenido que resolver dos retos importantes. Por un lado, capturar movimientos del actor para animar el avatar en tiempo real, en condiciones de espacio e iluminación limitados e incluyendo expresiones faciales y movimiento labial en función de la voz. Y por otro lado, integrar las imágenes reales con el actor virtual situando los objetos en su posición relativa correcta. Para la captura hemos analizado las tecnologías disponibles, como describiremos a continuación, optando por una combinación de ellas: un dispositivo Kinect para la captura de movimientos corporales del actor, un giróscopo para los movimientos de la cabeza y un WiiMote con algoritmos de control para gestionar las expresiones. Para la integración de imágenes, otro dispositivo Kinect nos permite obtener los mapas de profundidad y la imagen real para generar una imagen aumentada realista. Veamos cuáles son las tecnologías disponibles para la captura de movimiento.

1.1. Tecnologías de captura de movimiento con actores virtuales.

Los sistemas de captura de movimiento se han aplicado en numerosas ocasiones a la animación de personajes virtuales bien para conseguir secuencias de movimiento realistas de las diferentes partes del cuerpo y aplicarlas en animaciones, o bien para controlar actores virtuales que representan usuarios y cuyo movimiento se corresponde con el de éste en *tiempo real*. Estos sistemas, para ser eficientes, deben permitir calibrar de forma rápida a los usuarios, dar libertad de movimiento sin restricciones y conseguir un buen nivel de detalle. Tratando de cumplir estos requisitos encontramos diferentes tecnologías de sensorización para animar personajes virtuales, basadas en sistemas electro-mecánicos, acústicos [1], electro-magnéticos y ópticos.

Los *sistemas mecánicos* suelen disponer de una estructura exoesquelética con potenciómetros y se han utilizado en algunas aplicaciones de marionetas virtuales [2]. Tienen el inconveniente del cableado y la necesidad de calibración. Respecto a los sistemas *con sensores electromagnéticos* que constan de un emisor central y un conjunto de receptores colocados en las diferentes partes del cuerpo a capturar [3, 4, 5], pierden precisión conforme el usuario se aleja de la base emisora.

Por último, los *sistemas ópticos* se han convertido en los más populares debido a la facilidad de movimiento que ofrecen para los usuarios ya que no requieren de cableado. Estos sistemas frecuentemente hacen uso de marcadores (LEDS u otro tipo de materiales reflectantes) para detectar posición y movimientos por medio de dos o más cámaras, si bien en ocasiones hay problemas de oclusión de marcadores que hay que solucionar con un post-procesado. Comparando la posición de las imágenes en las diferentes cámaras se puede calcular la posición 3-D de cada marcador, obteniendo una lista de ellos y sus posiciones en el espacio.

También existen sistemas de detección de movimiento ópticos sin marcadores que utilizan algoritmos de reconocimiento de imágenes [6,7,8,9] para detectar posiciones o movimientos del cuerpo. Sin embargo, estos sistemas no suelen proporcionar una gran precisión y se usan principalmente en la detección de movimientos sencillos de algunas partes del cuerpo.

1.2 Realidad Aumentada y captura de movimiento.

Aunque tradicionalmente los sistemas de captura de movimiento para mover actores virtuales han sido utilizados en aplicaciones de Realidad Virtual o para la creación de animaciones, con el auge de las aplicaciones de R.A. encontramos el uso de estos sistemas para controlar al usuario dentro del mundo real en su interacción con los objetos virtuales en la escena aumentada.

Uno de los primeros ejemplos que encontramos es el sistema ALIVE [10] donde una cámara de vídeo capta una imagen de una persona con un objetivo doble, detectar los movimientos del usuario para generar acciones de los objetos virtuales, e integrar la imagen real en un entorno virtual. De esta forma se crea un sistema precursor de R.A. donde se integran imágenes de vídeo con objetos 3D (como un perro) que se activan mediante movimientos del usuario.

Existen también numerosos ejemplos de sistemas de R.A. que hacen uso de captura de movimiento con marcas (normalmente figuras geométricas simples) para detectar posiciones, sin embargo se trata casi siempre de la posición de unos pocos objetos. Estos sistemas no proporcionan la suficiente precisión cuando lo que queremos es mover un personaje virtual y todas sus articulaciones.

Otras aplicaciones de R.A. requieren capturar el movimiento fuera de laboratorios o entornos controlados, lo cual limita el tipo de dispositivos a utilizar, recurriendo por ejemplo a sistemas basados en ultrasonidos y sensores inerciales que detectan movimientos en contextos de uso muy amplios [11,12] o en sistemas ópticos con marcadores [13]. En estos sistemas el problema más importante es combinar la información real y virtual de manera que se produzcan oclusiones correctas en la integración.

2 Objetivos, requisitos y elección de tecnologías.

Como hemos indicado anteriormente, la idea principal de nuestra aplicación consiste en diseñar un personaje animado en tiempo real por un actor. Además, debe dar soporte a la voz del actor que simula la voz del personaje y la manipulación de objetos virtuales que reforzarán el discurso o la conversación. Estos objetos virtuales están integrados con el mundo real y son proyectados sobre una pantalla de grandes dimensiones a modo de *espejo aumentado*, permitiendo no solo la atracción de la audiencia sino la participación de toda ella en esta puesta en escena. Su aplicación práctica principal, el acto de inauguración del stand de la Diputación de Valencia en la Feria de Turismo de Madrid, el 19 de Enero de 2011.

Desde el punto de vista tecnológico teníamos que diseñar un sistema de R.A. que integrara imágenes de vídeo con imágenes virtuales, entre las cuales se encontraba un avatar controlado por medio de sistemas de captura de movimiento. Adicionalmente, existían una serie de requisitos derivados del contexto de uso del sistema, que iba a ser instalado en un stand de una feria para una ocasión puntual, con lo que se requería poca inversión en sistemas de captura de movimiento. Además, debíamos valorar la complejidad de transporte e instalación, así como las características de espacio reducido e iluminación no controlable que introducía restricciones a nuestro sistema.

Esto hizo necesario un estudio y selección de tecnologías que dieran soporte a todos estos requisitos. Revisando los sistemas de captura de movimiento descartamos los sistemas mecánicos y magnéticos por su mayor coste, complejidad en la instalación y calibración, no siendo adecuados para un solo uso en una demostración.

Respecto a los sistemas ópticos, quedaron también descartados los que utilizan los marcadores activos, ya que además de ser muy costosos por la calidad de las cámaras que utilizan, son difíciles de utilizar en entornos con espacio limitado e iluminación no controlada. Los sistemas que no utilizan marcadores, o los basados en marcas pasivas, podrían ser utilizados, ya que únicamente necesitan cámaras de bajo coste, sin embargo el gran inconveniente es la poca precisión en los movimientos.

Así, encontramos el sistema Kinect, aparecido recientemente para su uso en videojuegos. Este dispositivo, usado como método de captura de movimiento, requiere de un espacio mínimo de 2 metros y soporta muy diferentes condiciones de iluminación, no requiriendo que el usuario lleve puesto ningún elemento adicional o vestuario concreto. Gracias a la información de la cámara infrarroja que incorpora, es posible obtener mejor precisión para movimientos del cuerpo que los sistemas basados en marcas pasivas. Todo ello unido a un calibrado inmediato por medio de una postura estática del usuario predefinida, hace posible su uso en pocos segundos. Además, al incorporar una cámara RGB y un sensor de profundidad, Kinect nos proporciona también información de distancias y nos permite capturar imágenes, pudiendo ser usado para la generación del escenario aumentado final.

Decidimos por tanto explorar esta tecnología ya que además, cuando se desarrolló el proyecto, no encontramos referencias de otros sistemas que la utilizaran en este contexto, y pensamos que podía abrir un nuevo campo de uso. Recientemente han empezado a aparecer algunas aplicaciones basadas en esta tecnología [14]. Sin embargo este sistema nos creaba una serie de limitaciones que veremos con detalle cuáles fueron y cómo solventamos.

3 Descripción del sistema.

Dadas las características del sistema y las restricciones impuestas por el contexto de uso del mismo, es necesario establecer dos espacios de trabajo claramente diferenciados y separados físicamente, cada uno con un set de dispositivos. Por un lado tenemos el espacio orientado a la captura de movimiento del actor real, que permiten la comunicación con los espectadores, y la interacción del mismo con la aplicación controlando un conjunto de objetos animados. A este espacio lo denominamos **Escenario de Control** (Fig.1, derecha).

Por otro lado, tenemos el área donde está el público y los interlocutores que conversan con el personaje virtual. Este es el espacio que denominamos **Escenario Aumentado** (Fig.1, izda). En él se recoge información de la escena real y se aumenta con la información procedente del Escenario de Control, gestionándose toda la coordinación de información, coherencia visual y localización de los objetos en función de la escena real, proyectándose el resultado en una pantalla de grandes dimensiones a modo de *espejo aumentado*. Todo ello mediante una arquitectura basada en cliente-

servidor, que permite la comunicación de acciones, interacción con objetos y actualización de la escena en tiempo real. Veamos en detalle ambos escenarios.

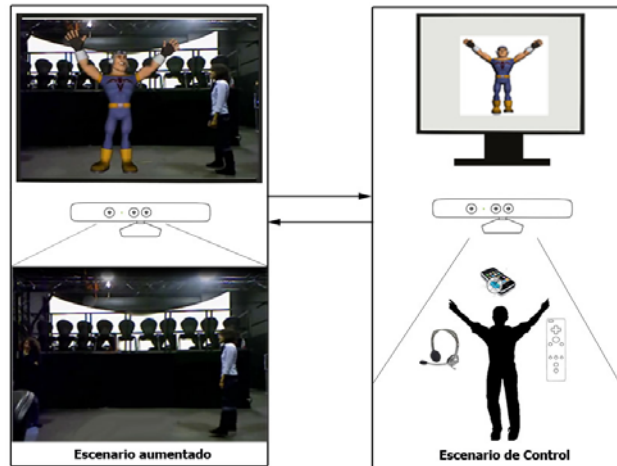


Fig. 1. Estructura general del sistema.

3.1 Escenario de Control: Sistema de captura de movimiento

La principal función de este escenario es el control del actor real en su interacción con la aplicación. Por un lado, su movimiento define las acciones del personaje virtual. Por otro lado, su interacción con otros dispositivos, tales como un mando de la Wii o un micrófono, producen tanto cambios en el personaje virtual como en la escena real. Toda la comunicación con los dispositivos utilizados la realiza una aplicación servidor que los transfiere a la aplicación cliente para su tratamiento.

3.2.1. Uso de Kinect y sus limitaciones.

El dispositivo principal usado en nuestro sistema para la captura de movimiento es el Kinect. Éste, a través de la SDK OpenNI (Open Natural Interface), proporciona información de las posiciones y orientaciones de un conjunto de articulaciones del esqueleto del actor real, que son traducidas por nuestra aplicación a posiciones y orientaciones finales para el esqueleto del personaje virtual.

A pesar de las ventajas ofrecidas por Kinect como sistema de captura de movimiento, también presenta diferentes limitaciones que afectan al realismo y naturalidad de los movimientos del personaje en nuestro sistema. Por un lado, el dispositivo no da información de la orientación de la cabeza en cada momento, ya que dicha articulación no está controlada por los drivers y librerías gratuitos actualmente disponibles. Esto supone una gran limitación, ya que es necesario que el personaje hable con los interlocutores del público y por tanto, oriente su cabeza hacia cada uno de ellos. Por otro lado, no controla el movimiento de los dedos de las manos, por lo que el personaje virtual presenta rigidez en las manos y poca expresividad al

interactuar con el público. Además, también es necesaria la animación facial y movilidad de la boca del personaje al hablar, para enriquecer la comunicación y la interacción con el público, así como la manipulación de diferentes objetos virtuales animados que complementen sus explicaciones, completando así la escena aumentada. Vamos a analizar las soluciones propuestas a cada problema planteado.

3.2.2. Soluciones aportadas a las limitaciones de Kinect

Para solucionar las limitaciones en la captura e interacción con la aplicación descritas anteriormente, completamos el sistema con la implementación de módulos auxiliares de control, algunos sincronizados con el movimiento del actor proporcionado por el dispositivo Kinect y otros basados en el uso de dispositivos auxiliares, también de bajo coste, que nos proporcionan la información adicional necesaria, manteniendo en todo momento la simplicidad y naturalidad de la interacción de cara al usuario (Fig.2).

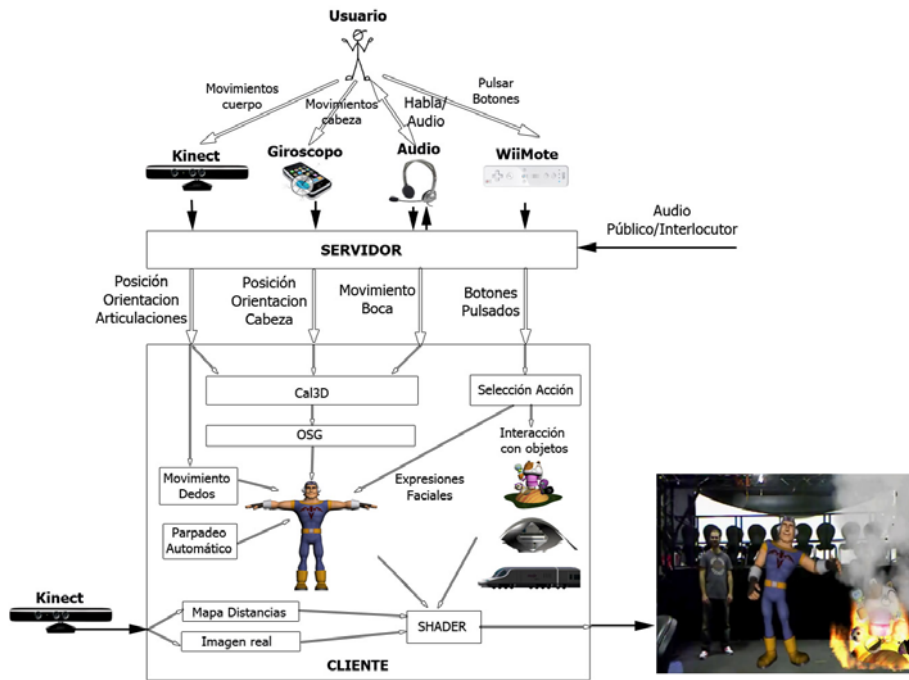


Fig. 2. Esquema la aplicación completa.

Para obtener el **movimiento de la cabeza** del actor real, información no proporcionada por el dispositivo Kinect, incorporamos al sistema un giróscopo colocado sobre la cabeza por medio de una gorra. Este elemento adicional no limitaba el movimiento del actor, ya que se comunica con la aplicación servidor por medio de

una conexión wireless a través de sockets, ofreciendo información de los movimientos de rotación realizados por la cabeza.

Por otro lado, **los dedos de las manos**, tan importantes en la expresión no verbal en la comunicación, no podían ser obtenidos por ningún sistema a no ser que se le colocara al actor real unos guantes de captura de movimiento, cuyo coste resultaba excesivo. Para resolverlo, implementamos un módulo de control automático de todos los dedos de las manos en función de la proximidad o lejanía de las mismas respecto al cuerpo, tal y como se mueven nuestros dedos en nuestras conversaciones cotidianas. Establecimos un algoritmo que calculando dicha distancia, determinara la flexión de cada falange de los dedos de cada mano, de manera que cuanto más cerca tenía el actor las manos del cuerpo, más cerrados estaban los dedos del avatar y al alejarse, se iban abriendo en consecuencia. Con ello, eliminamos la rigidez de los dedos del avatar y mejoramos la expresividad de los mismos en las conversaciones.

Dada la importancia de las **expresiones faciales y movilidad labial** para mejorar la comunicación no verbal entre las personas, es necesario incorporar al avatar ambos elementos. Al no disponer de información del dialogo que iban a mantener interlocutores y personaje, ni los cambios de estado de ánimo que requeriría el mismo en cada momento, se optó por integrar un módulo de control de la cara del avatar en tiempo real a partir de la información procedente del actor real. Por un lado, incorporamos un mando de la Wii al sistema como dispositivo de control de las expresiones faciales por parte del actor. Con él es posible modificar el estado de ánimo del avatar según cada instante de la conversación por medio de la pulsación de un conjunto de botones seleccionados, combinándose dichos movimientos con los del cuerpo y la cabeza. Por otro lado, el movimiento labial del personaje al hablar se incorporó por medio de un algoritmo de análisis de la amplitud de onda de la señal de audio procedente del actor real, capturada a través de un micrófono. Por medio de este algoritmo es posible calcular cuál es la amplitud tanto horizontal como vertical que debe tener la boca, y realizar una conversión de estos valores al sistema de referencia utilizado por los huesos que controlan la boca del avatar. El algoritmo además permite modificaciones en parámetros de control que facilitan la adaptación a la intensidad, volumen y velocidad del habla del actor.

Por último, la aplicación cliente, dispone de un módulo de generación automática del **parpadeo de los ojos del personaje**, que controla que, a pesar de las expresiones faciales activadas, el personaje siga realizando el parpadeo en un tiempo constante. Con todo ello, la expresividad del avatar queda completamente cubierta incorporando emociones en su rostro al hablar, parpadeo y movimiento de labios acorde con la conversación que se está manteniendo, lo que aporta naturalidad y humanidad a su interacción con el público.

La incorporación del mando de la Wii como dispositivo de control del avatar nos facilitó el uso del mismo como método de interacción con la escena real, para la incorporación de objetos virtuales a la misma en función de la conversación. La pulsación de ciertos botones produce la aparición de determinados objetos en miniatura en una de las manos del personaje, de manera que al soltarlo, se activa una animación que permite que el objeto vuele y se sitúe correctamente y al tamaño

adecuado en la escena real. Esto hace posible al público interactuar con los objetos virtuales como si estuvieran en el mundo real.

3.2.3. Estructura del actor virtual.

Como hemos indicado anteriormente, uno de los elementos virtuales que se incorpora a la escena real es un personaje controlado en tiempo real por un actor por medio de Kinect. Este dispositivo proporciona información sobre un conjunto determinado de articulaciones del esqueleto del usuario en su sistema de referencia concreto. Por ello, resulta imprescindible modelar el avatar y su esqueleto basándonos en la estructura articulada soportada por Kinect, por lo que, a partir de la malla del avatar, se ajusta todo su esqueleto, centrándonos en las articulaciones que necesitamos para desarrollar nuestro sistema de captura a medida, teniendo presente las especificaciones del dispositivo usado (Fig. 3a). Este esqueleto fue completado para dar soporte a los elementos adicionales incorporados por nuestro sistema: huesos para el control de la cabeza, para la movilidad facial (cejas, párpados, pómulos y labios) así como para la simulación del habla y el movimiento de los dedos de las manos (Fig. 3b).

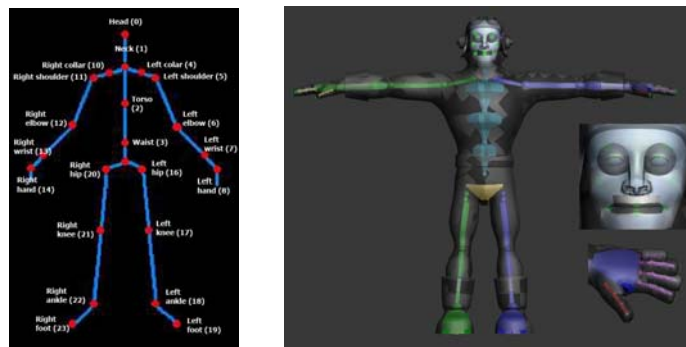


Fig. 3. a) Detalle del esqueleto soportado por Kinect. b) Avatar y esqueleto en nuestro sistema.

Para la representación gráfica del avatar, al igual que para los objetos virtuales que se integran en la escena, nuestro sistema hace uso de la librería gráfica de código abierto OpenSceneGraph (OSG), mientras que para almacenar la información relativa al avatar y su control a bajo nivel se utiliza la librería Cal3D [15]. Ya que esta última es independiente del motor gráfico final que se utilice, contamos con una capa de traducción de la información del avatar a nodos gráficos en OSG. Esto supone que, para animar el avatar en tiempo real, es necesaria la traducción de las posiciones y orientaciones entre los tres sistemas de referencia usados por cada capa de la aplicación: Kinect/Giroscopio – Cal3D – OSG, y su integración final en la escena real.

3.2 Escenario Aumentado: Sistema de visualización aumentada

El escenario aumentado es donde se realiza la interacción del público, con el actor y los objetos virtuales utilizando realidad aumentada. Para la captura de la imagen real y la información necesaria para su integración, utilizamos otro dispositivo Kinect situado estratégicamente en el stand. A través de una pantalla de grandes dimensiones, colocada junto al escenario, se ofrece la imagen capturada por Kinect aumentada con la información virtual, obteniendo como resultado un *espejo aumentado*. Kinect está equipado con dos cámaras, una convencional y otra infrarroja ambas con una resolución de 640x480 píxeles y una velocidad de captura de 30 imágenes por segundo, de manera que utilizando el SDK OpenNI, es posible obtener además del flujo de video convencional, un flujo de imágenes de profundidad. Mezclando esta información con los objetos virtuales, es posible calcular que parte de los objetos virtuales se encuentran delante y cuales detrás de los usuarios. Como resultado el usuario se ve en el *espejo aumentado* rodeado por los distintos objetos virtuales, pudiendo caminar alrededor de ellos. Para realizar una correcta mezcla de esta información de distancias se ha implementado un shader que escribe al mismo tiempo la imagen de vídeo en el buffer de color y calcula el buffer de profundidad (Z-Buffer [15]). Aplicando el shader a la escena virtual se obtiene una mezcla perfecta entre el mundo real y virtual, sin ralentizar el funcionamiento de la aplicación puesto que el shader se ejecuta directamente en la GPU de la tarjeta gráfica.



Fig. 4. Escena aumentada en dos escenarios diferentes aplicando el shader de oclusión.

Con todo ello, es posible incorporar objetos virtuales, como el personaje, así como efectos que modifican tanto los objetos como la escena real. Un ejemplo es la incorporación de fuego y humo basado en un sistema de partículas utilizando el `odekit osgParticle`, incluido en OSG, usado en nuestro caso para simular las fallas valencianas en la presentación, consiguiendo resultados muy realistas (ver Fig. 4).

4 Conclusiones

En el presente artículo hemos descrito un sistema de interacción integrando diferentes tecnologías en una aplicación de R.A. Gracias a la selección de dispositivos realizada y a los módulos de control desarrollados, el resultado ha sido un sistema que funciona

correctamente ante condiciones de espacio limitado e iluminación no controlada, con una gran sencillez de instalación y uso. Destacamos la estabilidad de la captura de movimiento, la expresividad del avatar integrado y la correcta oclusión en la visualización aumentada, obtenida aplicando las distancias calculadas por Kinect al proceso de dibujado de los objetos virtuales sobre la escena real.

En su presentación en Fitur'2011 el sistema fue utilizado dentro de un stand de la Diputación de Valencia. Se situó al actor real, que controlaba y daba vida al avatar, oculto al público, y en el stand, una pantalla gigante (*espejo aumentado*) permitía visualizar imágenes del público y del presentador del acto, combinadas con elementos virtuales, como el avatar con el que se conversaba y otros objetos. Durante esta conversación, que versaba sobre recursos de la Comunitat Valenciana, aparecían diferentes animaciones haciendo referencia al AVE, a las fallas o la paella valenciana.

Como trabajo futuro nos planteamos probar el nuevo SDK lanzado por Microsoft y comparar resultados principalmente en lo que respecta a la precisión de los datos del tracking de las articulaciones para evitar filtrados posteriores. Otra posible mejora será conseguir mejor movimientos del avatar utilizando cinemática inversa.

5 Bibliografía

1. Sugimoto, M. Tulathimutte, K., Ito, T., Sato, T., Hashizume, H. An Ultrasonic 3D Positioning System Using a Single Compact Receiver Unit. Proc. of LoCA'09. (2009)
2. Walters, G. The story of Waldo C. Graphic. Course Notes: 3D Character Animation by Computer, ACM SIGGRAPH '89, Boston, (1989)
3. Slyper, R., Hodgins, K. Action Capture with Accelerometers. Eurographics / ACM SIGGRAPH Symposium on Computer Animation (2008)
4. Jessel, J-P., Jaspert, C., Flores, J-J. Computer Animation and Virtual Reality for Live Art Performance. Virtual storytelling: using V.R. technologies for storytelling. (2001).
5. Huang, Y. Kallman, M. Interactive Demonstration of Pointing Gestures for Virtual Trainers. LNCS 5611. pp178-187. (2009)
6. Wren, R., Azarbayejani, A., Darrell, T. and Pentland, A. P. . Pfunder: real-time tracking of the human body, Trans. Pattern Anal. Mach. Intelligence 19(7), 780-785 (1997)
7. Pentland, A., et al.. Real-time 3D Motion Capture. In proceedings Microsoft Research (1998)
8. Plankers, R. Fua, P., D'Apuzzo, N. Automated Body Modeling from Video Sequences. IEEE International Workshop on modeling people at ICCV'99, Greece (1999)
9. Pinhanez, C., Bobick, a. Using Computer Vision to Control a Reactive Computer Graphics Character in a Theater Play. Proc. of ICVS'99 (1999)
10. Maes, P., Darrell, T., Blumberg, B. Pentland, A.. The ALIVE system: wireless, full-body interaction with autonomous agents. Multimedia Systems 5: 105-112 (1997)
11. Foxlin, E., Harrington, M., Pfeifer, G. Constellation: A Wide-Range Wireless Tracking System for Augmented Reality and virtual set applications. Proc. of SIGGRAPH'98. (1998).
12. Vlasic D. et al. Practical motion capture in everyday surroundings. ACM Transactions on Graphics, Vol. 26,3.(2007)
13. Dorfmüller, K. Robust tracking for augmented reality using retroreflective markers. Computer & Graphics. Volume, 23, 6, 795-800 (1999).
14. Kimber, D. Vaughan, J., Rieffel, E. Augmented Perception through Mirror Worlds. Conference AH'11 (2011).
15. Breen, D. E., Whitaker, R. T., Rose, E. and Tuceryan, M. Interactive Occlusion and Automatic Object Placement for Augmented Reality. Computer Graphics Forum, 15: 11-22. (1996)