

# Interacción con Entornos Virtuales en teléfonos móviles

Jesús Gimeno , Inmaculada Coma, Marcos Fernández, Pedro Morillo, Lucía Vera.

Instituto de Robótica, Universitat de Valencia

{Jesus.Gimeno, Inmaculada.Coma, Marcos.Fernandez, Pedro.Morillo, Lucia.Vera}@uv.es

**Abstract.** En este trabajo presentamos un sistema para la navegación en entornos virtuales 3D sobre teléfonos móviles. Tradicionalmente la interacción con el teléfono móvil se ha realizado utilizando el teclado como dispositivo de interacción. Sin embargo, gracias a las capacidades de los nuevos terminales es posible aprovechar dispositivos como la cámara incorporada o los acelerómetros para implementar interacciones basadas en movimientos del usuario. En el presente trabajo hemos tratado de crear un mecanismo de interacción con entornos virtuales que sea adaptable a las tecnologías incorporadas en el teléfono, de forma que se pueda interaccionar bien interacción mediante las teclas, o bien capturar los movimientos del usuario haciendo uso de los acelerómetros o las imágenes capturadas por la cámara de vídeo. Para estimar el movimiento a partir de las imágenes capturadas se ha utilizado un algoritmo diferencial. Por último se ha diseñado un entorno 3D que permite el acceso a contenidos Web y multimedia.<sup>1</sup>

**Keywords:** Navegación 3D, Teléfonos móviles, Captura basada en vídeo.

## 1 Introducción

El avance en las tecnologías hardware en los teléfonos móviles de última generación ha permitido que aplicaciones de Realidad Virtual o Realidad Aumentada puedan ser ejecutadas sobre este tipo de dispositivos. Así, es posible superponer información sintética sobre imágenes reales [9] o utilizar juegos 3D.

Para la interacción con los teléfonos móviles tradicionalmente se ha utilizado el teclado y el joystick o los cursores, lo cual limita bastante el tipo de interacción y la hace poco adecuada para el movimiento en entornos virtuales.

La introducción en los teléfonos móviles de alta gama de dispositivos como acelerómetros, pantallas multi-táctiles, y pantallas de mayores dimensiones ha mejorado enormemente la experiencia del usuario. El uso de acelerómetros permite implementar técnicas de interacción basadas en movimientos del usuario, que suelen ser más adecuados que el uso del teclado para la navegación en entornos virtuales. Sin embargo, el uso de acelerómetros todavía está limitado a algunos teléfonos móviles de alta gama.

---

<sup>1</sup> Esta investigación ha sido parcialmente financiada por la compañía Orange R&D España y el plan nacional de I+D a través del proyecto DPI 2006-14908-C02-02

En este trabajo proponemos un sistema de interacción 3D que puede utilizar diferentes dispositivos de entrada presentes en los teléfonos móviles. De esta manera el movimiento del usuario puede obtenerse mediante tres mecanismos de entrada diferentes: la interacción con el teclado, la interacción a través de movimientos detectados por los acelerómetros (si están presentes en el terminal) o la interacción a través de movimientos detectados haciendo uso de las imágenes de vídeo capturadas de la cámara.

El propósito de este sistema de interacción 3D que sirve para múltiples dispositivos es proporcionar flexibilidad al sistema, ya que es posible adaptarse a las capacidades hardware de teléfono sobre el que funciona. Así, si el teléfono dispone de cámara o acelerómetros, éstos serán utilizados como mecanismos de interacción, pero también podremos interactuar con el entorno 3D en aquellos terminales en los que solo dispongamos de teclado.

En el presente trabajo, además del sistema de interacción, ha sido diseñado un entorno virtual consistente en una tienda donde el usuario puede acceder a diferentes servicios ofrecidos por el operador Orange, que permita el acceso a contenidos Web, manejar contenidos multimedia, que incluya gráficos 3D en tiempo real y un interfaz accesible desde diferentes plataformas. Entre los servicios disponibles en el entorno virtual se incluye el acceso a noticias, canales de televisión, ofertas de productos y acceso a la Web. Los resultados obtenidos en la evaluación del sistema muestran que la aproximación que presentamos puede ser validada como un interfaz 3D multiplataforma para teléfonos móviles.

El presente artículo está organizado de la siguiente forma: en el apartado 2 se realiza una revisión de los trabajos relaciones con las técnicas de interacción sobre móviles así como los problemas de las soluciones existentes. En el apartado 3 se describe el sistema completo, la arquitectura software de la aplicación, el motor 3D utilizado, los escenarios creados y las técnicas de interacción implementadas con los distintos dispositivos, y por último en el apartado 4 se exponen las conclusiones.

## **2 Trabajos previos.**

Con el objetivo de mejorar la interacción del usuario sobre dispositivos móviles se han desarrollado un cierto número de trabajos en los que se proponen nuevas técnicas de interacción. En este sentido encontramos trabajos que hacen uso de diferentes tecnologías como dispositivos de entrada.

Por ejemplo, la introducción de sensores táctiles [7] permite al usuario una interacción directa con los interfaces mediante la selección de opciones. En esta línea algunos dispositivos móviles utilizan el reconocimiento de gestos como el HTC TouchFLO3D que permite la detección de cuatro movimientos, o la pantalla multitáctil de Iphone que permite operaciones como zoom o rotación de imágenes.

Los sensores de proximidad y los acelerómetros [9] también son utilizados como dispositivos de entrada. Hinckley [7] utiliza los acelerómetros para la detección de la visualización horizontal o vertical en la pantalla, y Wigdor [16] hace uso de los mismos para la introducción de textos o el desplazamiento mediante scroll a través de

documentos. Sin embargo, el principal inconveniente de dichas tecnologías es que de momento sólo las encontramos presentes en teléfonos de alta gama.

Otra de las tecnologías utilizadas como dispositivo de entrada es la videocámara, presente en una gran cantidad de teléfonos móviles. El desarrollo de diferentes algoritmos y técnicas para la detección de movimientos basados en las imágenes de vídeo proporcionadas por la cámara ha posibilitado nuevas técnicas de interacción.

Numerosos investigadores y compañías utilizan sistemas de reconocimiento de códigos mediante la cámara [11,13]. El reconocimiento de estos códigos suele ir asociados a diferentes acciones. Rohs[12] y Wagner[14] han desarrollado un sistema para detección de códigos visuales 2D. Mohring [10] propone un sistema para el seguimiento de formas coloreadas en 3D. Estas soluciones son muy utilizadas en los sistemas de Realidad Aumentada para teléfonos móviles donde se crean aplicaciones que utilizan las imágenes de la cámara y superponen sobre ella algún tipo de geometría 3D. El uso de marcadores permite la obtención de movimientos absolutos del teléfono, sin embargo, restringe la utilidad del sistema debido a la necesidad del uso de marcadores.

Una aproximación diferente es el uso de algoritmos de detección que no necesitan marcadores pero que obtienen movimientos relativos. Algunos investigadores han implementado algoritmos basados en la detección de características, que buscan en las imágenes bordes o esquinas tratando de estimar con ellas la dirección y magnitud del movimiento entre diferentes imágenes. Un ejemplo de este tipo de aplicaciones nos lo da Haro [1][4], que utiliza un algoritmo basado en el seguimiento de formas para detectar el movimiento, y lo utiliza en diferentes aplicaciones incluyendo un visor de documentos, un visor de fotos, así como juegos 2D. También implementa un juego 3D basado en Quake III donde no se utilizan ni texturas ni cálculos de iluminación obteniéndose una tasa de entre 3 y 10 frames por segundo. Hannuksela [6] incluye el uso de filtros de kalman en sus algoritmos de detección de características obteniendo técnicas más eficientes. También existen ejemplos de juegos de Realidad Aumentada basados en la detección de color [5]. En ocasiones se utilizan estos algoritmos de detección para navegar en entornos virtuales y seleccionar objetos, como es el caso de Hwang [8], aunque en este caso el dispositivo móvil es un PC.

La principal desventaja de los algoritmos de detección basados en características es que cuando se enfrentan a grandes diferencias de iluminación, sombras, o entornos planos en los que no hay características relevantes no funcionan correctamente, además del alto coste computacional que suele implicar el algoritmo de detección.

Otro tipo de sistemas que no utilizan marcadores son aquellos basados en algoritmos que buscan diferencias entre imágenes y aplican la estimación de movimiento, similares a las técnicas empleadas en los codificadores de vídeo MPEG. Un trabajo con algoritmos de este tipo Tinymotion [15], que utiliza un algoritmo diferencial compuesto de cuatro pasos: conversión de espacio de colores, muestreo, estimación de movimiento y post procesado [15]. Este software ha sido evaluado para ser utilizado como puntero así como en operaciones de selección de menús, entrada de texto o juegos 2D como el tetris.

En nuestro trabajo hemos hecho uso de algoritmos diferenciales principalmente por dos razones. Por un lado, el uso de marcadores no es adecuado en aplicaciones donde el entorno no está controlado, y por otra parte los algoritmos diferenciales trabajan mejor que aquellos basados en características cuando nos encontramos con fondos

planos. Sin embargo, como explicaremos más adelante, hemos introducido algunas mejoras en el algoritmo para evitar algunos problemas como el auto-balance de blancos en las cámaras.

### 3 Descripción del sistema

Empezaremos describiendo la arquitectura del sistema y los componentes que lo conforman. El sistema está compuesto por un conjunto de componentes distribuidos entre el servidor Web y el teléfono móvil. En la figura 1 se muestra la relación entre ellos y pasamos a describir los más importantes a continuación.

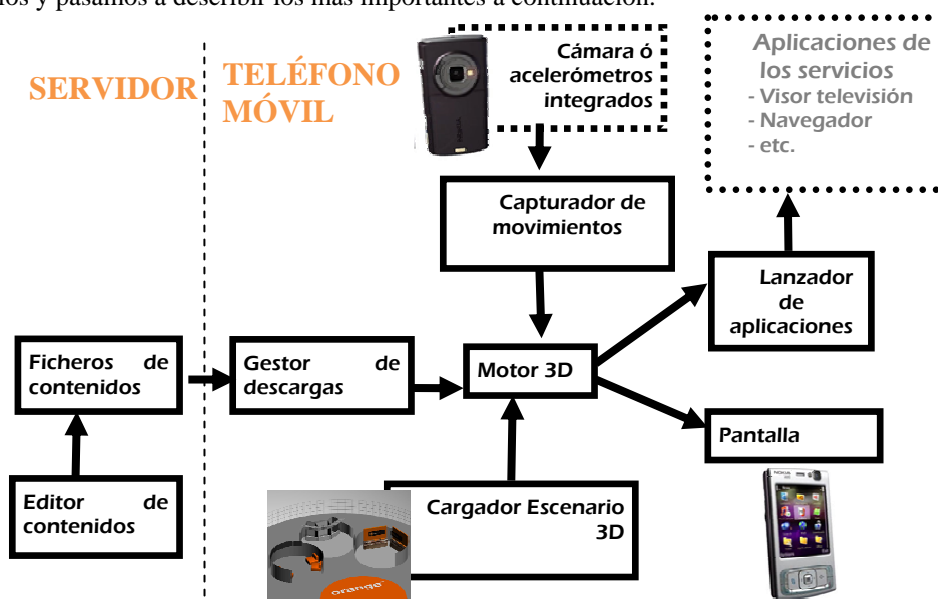


Figura 1. Diagrama de la arquitectura y componentes del sistema.

#### 3.1. Edición de contenidos en el servidor.

Por un lado tenemos en el servidor un sistema editor de contenidos que nos permite generar ficheros de contenidos. El editor es un software que permite al administrador de la aplicación generar los ficheros de contenidos del portal. Para ello se ha implementado con Windows Forms un interfaz gráfico a través del cual se configuran los enlaces que aparecerán posteriormente en el portal. Este módulo es el intérprete entre el usuario administrador y el portal, permitiendo al usuario actualizar los contenidos mediante un interfaz gráfico. Una vez seleccionados los contenidos, se generan los ficheros de automáticamente.

En la aplicación piloto desarrollada para Orange (Orange 3D World) se ha definido un formato XML que contiene la descripción del portal y los diferentes enlaces que

aparecen. El portal tiene información clasificada en cuatro categorías: televisión, noticias, tienda y publicidad, cuyos contenidos se almacenan en el fichero de contenidos y pueden ser modificados desde el editor.

### **3.3. Gestión de contenidos desde el teléfono móvil.**

La aplicación portal instalada en el teléfono móvil está compuesta por los siguientes módulos: el gestor de descargas que recoge e interpreta los ficheros de contenidos, el motor 3D que visualiza la información a partir del escenario 3D y los movimientos del usuario, el capturador de movimientos que procesa la información de los diferentes dispositivos del teléfono para calcular el movimiento del usuario y el lanzador de aplicaciones que lanza aplicaciones externas. Veamos cómo se han implementado estos módulos, y en el siguiente apartado describiremos con más detalle el sistema de captura de movimientos.

#### **3.3.1. Cargador escenario 3D.**

Para el desarrollo de los modelos 3D se ha hecho uso de la herramienta de diseño 3D Studio Max, de Autodesk, que es uno de los más utilizados en este tipo de tareas y permite el uso de plugins. Estos plugins son creados previamente utilizando C++ y se incorporan al programa como una librería dinámica DLL. Para poder importar los modelos creados en el sistema se ha creado un módulo compuesto de un exportador y un cargador. El exportador se instala como un plugin en el 3D Studio Max, y añade la capacidad al editor de guardar los modelos en el formato .mod, implementado por los desarrolladores. Una vez creado el fichero dentro de la aplicación del portal el cargador traduce este fichero a instrucciones OpenGL que serán las utilizadas por el Motor 3D para dibujar el entorno 3D.

El exportado de los modelos 3D se ha realizado teniendo en cuenta que van a ser utilizados en un teléfono móvil. Aunque estos dispositivos tienen cada vez más capacidad de procesamiento, siguen siendo muy inferiores a un computador dotado de una tarjeta gráfica especializada. Por esto al guardar el fichero se ha realizado una optimización del modelo, tanto a nivel de texturas para ahorrar espacio en memoria, como a nivel de geometría para optimizar el tiempo de procesado, y de nuevo, la memoria necesaria para el mismo.

#### **3.3.2. Motor 3D.**

Este módulo es el encargado de todo el manejo de la escena 3D. En este módulo se implementan todas las funcionalidades necesarias para, partiendo de un modelo 3D y la posición del punto de vista del usuario, obtener una imagen 2D que se mostrará en la pantalla del teléfono móvil.

Una de las características más interesantes de este módulo es, que aunque utiliza el estándar OpenGL (versión simplificada OpenGL ES), permite manejar estructuras más complejas sin necesidad de definir las en forma de triángulos. Con este manejo de la información gráfica a alto nivel, es posible crear un grafo de escena, con el que mantener internamente una estructura coherente de todos los elementos a visualizar (el entorno contiene zonas, la zona de televisión contiene televisiones, etc.). Este

grafo de escena, combinado con la posibilidad de cargar modelos complejos generados con una herramienta de diseño comercial (3D Studio Max), permite obtener un desarrollo de aplicaciones mucho más rápido y eficiente.

### **3.4. Escenarios 3D**

Para el prototipo desarrollado se han diseñado una serie de escenarios 3D que simulan una tienda virtual en la que el usuario puede acceder a diferentes servicios. Durante el diseño se ha pensado en crear un entorno adecuado para las capacidades de interacción que se ofrecerán. Estas capacidades, tienen como entrada principal desplazamientos laterales, bien capturados por la cámara o bien por los acelerómetros (si el teléfono dispone de ellos). Bajo esta premisa la información se presentará principalmente en formato horizontal, con el fin de utilizar estos desplazamientos capturados para navegar fácilmente por las distintas partes de este entorno tridimensional.

El entorno virtual consiste en una serie de salas integradas dentro de lo que se ha llamado el 3D Orange World como se muestra en la figura 2. Este entorno 3D presenta las diferentes salas al usuario a su alrededor, desde una posición central. Con esta disposición, utilizando los desplazamientos laterales el usuario puede dirigir su punto de vista a la sala que desea visitar. Una de las decisiones de diseño adoptadas, es mostrar salas con planta circular. Estas salas circulares al tener una planta común podrán actualizarse fácilmente si se cambian los servicios ofrecidos en esta aplicación, sin modificar la estética general de la aplicación. De esta forma aunque cambien los servicios, el usuario sigue teniendo la sensación de que está manejando la aplicación que ya conoce, y sabe manejar.

Las salas diseñadas han sido:

- Sala de televisión, donde se presentan diferentes pantallas de televisión que simbolizan los distintos contenidos multimedia disponibles. Para ambientar esta sala se han utilizado elementos típicos de una sala de proyección como los sillones, o los televisores.
- Sala de noticias. En esta sala se presenta un kiosco de periódicos sobre el que se colocan las diferentes fuentes de noticias en disposición horizontal, en este caso diferentes periódicos. El usuario podrá desplazar lateralmente las opciones para elegir la deseada. Para ambientar esta sala se ha utilizado un kiosco de periódicos. Viendo este kiosco el usuario fácilmente puede entender que va a encontrar servicios relacionados con las noticias y los periódicos.
- Sala de tienda. En esta sala se dispone de un el mostrador de una tienda sobre el que se deslizan los anuncios de productos ofrecidos en la tienda. Con los desplazamientos horizontales los anuncios se van trasladando a la izquierda o a la derecha, de forma circular, al igual que en las secciones anteriores.

### **3.5. Interacción con el sistema.**

Uno de los retos del presente trabajo ha sido desarrollar una técnica de interacción con el entorno virtual que pudiera adaptarse a los diferentes dispositivos móviles. Para

ello hemos implementado un sistema de navegación que puede utilizar tres dispositivos de entrada de eventos: el teclado, los acelerómetros o la cámara de vídeo. Las tecnologías presentes en el teléfono móvil serán las que determinen el tipo de dispositivo de interacción que se utilice, y aunque los acelerómetros y la cámara proporcionarán una mejor experiencia de usuario ya que permiten la interacción con movimientos, será necesario el uso del teclado en determinados dispositivos. La interacción con el entorno virtual ha sido diseñada de tal forma que para cualquiera de los dispositivos de entrada utilizados el sistema funcione de la misma manera.

Respecto a la interacción basada en el teclado se trata del mecanismo más sencillo y aunque consideramos que el menos natural para moverse por entornos virtuales ha sido necesario desarrollarlo, tal y como hemos comentado, para que el sistema se adapte a todo tipo de dispositivos. Así, se recogen cuatro eventos correspondientes a los movimientos arriba, abajo, izquierda y derecha. Estos cuatro movimientos serán iguales con los acelerómetros y con la cámara de forma que el sistema de interacción será el mismo con cualquiera de los dispositivos.

#### **Captura basada en acelerómetros.**

Este tipo de captura consiste en utilizar los acelerómetros incorporados en el teléfono móvil, en el caso de que los lleve incorporados. Los acelerómetros miden las aceleraciones en los tres ejes, por tanto para obtener el movimiento en cada uno de los sentidos observados (arriba, abajo, izquierda y derecha) sólo es necesario comprobar las aceleraciones sufridas en los ejes correspondientes.

Este tipo de captura calcula directamente los movimientos del teléfono y no requiere un gran procesamiento, simplemente un pequeño filtrado de las señales de los acelerómetros para evitar convertir cualquier pequeño movimiento del usuario en movimientos en el entorno virtual. Para ello, el filtro utilizado establece un nivel por debajo del cual los movimientos realizados no son tenidos en cuenta. Este nivel ha sido calibrado empíricamente con diferentes usuarios.

Además, dado que no necesitamos posiciones absolutas sino movimientos relativos, una vez hemos detectado un movimiento en los siguientes se supone que el usuario está volviendo a su posición inicial, descartándose como movimientos los realizados en la dirección opuesta, que son tomados como retorno a la posición inicial.

Este filtrado de datos elimina uno de los problemas que aparece a menudo con el uso de los acelerómetros como es el ‘error de deriva’. Este error es la suma de la acumulación de errores cuando se intenta calcular un movimiento absoluto a partir de datos relativos.

#### **Captura basada en cámara utilizando un algoritmo diferencial.**

Este tipo de captura se basa en utilizar técnicas de visión por computador para detectar el movimiento del teléfono a partir de las imágenes capturadas por la cámara integrada, y transformar estos movimientos en eventos de entrada.

Tal y como hemos mencionado en la revisión, hemos optado por la detección haciendo uso de un algoritmo diferencial. Comparando las diferentes técnicas, hemos considerado ésta como la más adecuada para nuestra aplicación, ya que la detección basada en marcadores no puede ser utilizada en cualquier contexto, y los algoritmos basados en características ofrecen peores resultados.

Nuestro algoritmo es similar al utilizado en TinyMotion [15], pero hemos introducido algunas modificaciones para resolver los problemas encontrados y

mejorar la eficiencia general del sistema. Así, la aproximación implementada consta de tres pasos: binarización, muestreo y estimación de movimiento. Como no utilizamos medidas absolutas de movimiento no será necesario realizar un post-procesado de los datos.

Respecto al primer paso, la conversión de espacio de colores, inicialmente se implementó utilizando una conversión a escala de grises. De esta forma una imagen RGB de 24 bits es convertida en otra de escala de grises de 8 bits. Sin embargo, con la detección de movimientos usando escala de grises se encontró un serio problema: todos los teléfonos utilizados incluían balance de blancos automático y no podía desactivarse esta opción. Así, los grises obtenidos se modifican rápidamente con los cambios de luz debido al balance de blancos causando muchos errores de detección.

Para resolver este problema hemos utilizado la *binarización* en lugar de la escala de grises, utilizando un umbral variable en función del valor medio de los colores de los píxeles, de forma que el balance de blancos también modificará este umbral y los píxeles por encima y por debajo de él permanecerán fijos. Además, la binarización nos permitirá utilizar menos espacio en memoria.

Para reducir el tiempo de computación y la memoria utilizada para la estimación de movimiento se ha *muestreado en una malla*, tomando una ventana de 5x5 píxeles y obteniendo una malla de 25x20 celdas.

El último paso del proceso es la *estimación de movimiento*, que consiste en utilizar la matriz de 25x20 obtenida como un macro-bloque para realizar la estimación de movimiento. Así, se comparan dos imágenes consecutivas y se buscan los movimientos de píxeles en cuatro direcciones utilizando un rango de 5 píxeles para la comparación. Hemos utilizado un rango de búsqueda mayor que en otras aproximaciones (5 píxeles en lugar de 3 [5]) y con ello se obtienen mejores resultados en la estimación. Esto ha sido posible gracias en parte a la binarización ya que las comparaciones de 1 bit son más rápidas que las de 1 byte. Con ello, hemos obtenido una mejor detección de los movimientos rápidos del usuario.

Con este sistema somos capaces de detectar los movimientos del usuario en cualquier dirección, pudiendo utilizarse como eventos de entrada del sistema de navegación. Para crear una interacción consistente con los movimientos realizados con teclas y acelerómetros se han detectado también cuatro movimientos, y usando el mismo criterio los pequeños movimientos en una dirección por debajo de un umbral han sido despreciados.

#### **Navegación por el entorno.**

Utilizando cualquiera de los interfaces de entrada descritos (teclado, acelerómetros o cámara) el resultado es un sistema que permite navegar por el portal mediante cuatro movimientos: izquierda, derecha, arriba y abajo. Dependiendo del contexto en el que se encuentre el usuario los movimientos serán utilizados para realizar un tipo de acción u otro. Así, los movimientos laterales son utilizados para modificar el punto de vista dentro del entorno virtual, y entrar dentro de una sección u otra, o ya dentro de una sección pueden servir para cambiar la selección del servicio al que accedemos. Con los movimientos arriba y abajo podemos movernos hacia delante, para seleccionar una opción del interfaz o podemos movernos hacia atrás.





**Figura 2.** Diferentes imágenes del entorno 3D y el interfaz con un Nokia N95 y con un Iphone.

## 4 Conclusiones

En el presente trabajo hemos propuesto un sistema de navegación en teléfonos móviles que puede utilizar como interfaz de entrada los diferentes dispositivos presentes en un teléfono: teclas, acelerómetro o video cámara. Con ello se consigue un sistema de interacción basado en movimientos en aquellos dispositivos que alta gama que disponen de cámara y/o acelerómetro, pero también un entorno que se puede adaptar a las capacidades de los teléfonos móviles de gama más baja.

El uso de las imágenes de la cámara como dispositivo de entrada ha sido implementado adaptando un algoritmo diferencial para la detección de movimiento en imágenes de vídeo, mejorando algunos problemas surgidos al aplicar el algoritmo tal y como ha sido utilizado en otros trabajos.

La eficiencia del sistema ha sido testada en diferentes condiciones. El frame rate del sistema 3D ha sido medido, visualizándose un entorno de aproximadamente 22.000 vértices, utilizando texturas, a una tasa de 22 fps en un teléfono Nokia N95 cuando se interactúa mediante teclas o acelerómetros. Combinando el algoritmo de captura con la visualización la tasa obtenida ha sido de 15 fps.

Respecto a la tasa de adquisición de la cámara mientras el algoritmo de detección de movimiento está siendo ejecutado se obtiene una tasa de 30 fps. Por lo tanto,

podemos decir que el procesamiento del algoritmo se produce en tiempo real, lo cual es un resultado mejor que el obtenido en otros trabajos [15].

## Bibliografía.

1. Capin, T., Haro, A., Setlur, V., Wilkinson, S.,. Camera-based Virtual Environment Interaction on Mobile Devices. In Lecture Notes in Computer Science 4263, pages 765-773 October (2006)
2. Coma, R. Rodriguez, M. Fernández, E. Martínez, P. Caselles. Artgraph: un entorno integrado de desarrollo y ejecución de aplicaciones 3D tiempo real. In *Proc. of CEIG'98*. Spain, (1998)
3. Davis, J., Bobick, A.. The representation and recognition of human movement using temporal templates. In *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Pages 928-934, Puerto Rico (1997)
4. Haro, A., Mori, K., Capin, T., Wilkinson, S.. Mobile Camera-based User Interaction. *Proc. of IEEE International Conference on Computer Vision Workshop on Human-Computer Interaction* , pages 79-89, China (2005)
5. Hakkarainen, M., Woodward, C.. Symball- Camera Driven Table Tennis for Mobile Phones. ACM International Conference Proceeding Series; Vol. 265, pages 391-392. Spain, (2005)
6. Hannuksela, J., Sangi, P., Heikkilä, J.. Vision-based motion estimation for interaction with mobile devices. *Computer Vision and Image Understanding*. Volume 108, issue 1-2 pages 188-195. October (2007)
7. Hinckley, K., Pierce, J. , Sinclair, M., Horvitz, E.. Sensing techniques for mobile interaction. *Proc. of the 13<sup>th</sup> annual ACM symposium on User interface software and technology* , pages 91-100. USA (2000)
8. Hwang, J., Kim, G. J., Kim, N.. Camera based Relative Motion Tracking for Hand-held Virtual Reality. *Proc. of NICOGRAPH International 200,6* Seoul (2006)
9. Kähäri, M., Murphy, D. J.. MARA- Sensor Based Augmented Reality System for Mobile Imaging Device. In *Proc. of Fifth IEEE ISMAR*, USA (2006)
10. Möhring, M. , Lessig, C. , Bimber, O.. Video see-through AR on Consumer Cell-Phones. In *Proc. of the 3rd IEEE/ACM ISMAR* , pages 252-253. USA (2004)
11. Rekimoto, J., Ayatsuka, Y. CyberCode: designing augmented reality environments with visual tags. In *Proc. of DARE 2000* ,pages 1-10, Denmark (2000)
12. Rohs, M., Gfeller, B. Using Camera-Equipped Mobile Phones for Interacting with Real-World Objects. In *Advances in Pervasive Computing, Austrian Computer Society (OCG)*, pp. 265-271, (2004)
13. SEMACODE <http://semacode.com/>
14. Wagner, D. , Tobias, L. , Dieter, S.. Robust and Unobtrusive Marker Tracking on Mobile Phones. *Proc. of the 7th ISMAR*, pp.121-124, UK (2008)
15. Wang, J., Canny, J. TinyMotion: Camera Phone Based Interaction Methods. *Proc. of Conference on Human Factors in Computing Systems*. pp. 339-344, Canada (2006)
16. Wigdor, D. , Balakrishnan, R. TiltText: Using tilt for text input to mobile phones. *Proc. of UIST*. pages 81-90. Canada (2003)