

Capacitación en el uso de Equipos Médicos utilizando Herramientas de Realidad Aumentada en Dispositivos Móviles

Pablo A. Roncagliolo y Agustín J. González

*Departamento de Ingeniería Electrónica Universidad Técnica Federico Santa María, Chile
pronca@atmlab.utfsm.cl*

Resumen— Durante los últimos años los equipos móviles, celulares y PDAs, poseen cada vez mayores capacidades y flexibilidad en su programación, e incorporan diversos dispositivos periféricos: canales de comunicación inalámbrica y características multimedia. Esta gran cantidad de tecnologías transforman a los dispositivos móviles, de nueva generación, en excelentes herramientas para la captura y el manejo de información móvil, pues su portabilidad, creciente capacidad de procesamiento, autonomía, y gran variedad de tecnologías de adquisición y comunicación, abren las puertas a un campo ilimitado de aplicaciones. Una de estas posibilidades es el desarrollo de aplicaciones de Realidad Aumentada sobre dispositivos móviles, que consiste en la superposición referenciada de leyendas u objetos virtuales sobre la imagen capturada en tiempo real por la cámara. El presente trabajo aborda la aplicación de esta tecnología a la capacitación en el uso de equipos médicos, a través del uso de celulares o PDAs. Mediante algoritmos de procesamiento de imágenes la aplicación desarrollada permite decodificar el número del equipo y la distancia del dispositivo móvil. De esta forma, por ejemplo, se superponen punteros, indicaciones y textos, asociados a los controles del equipo médico en cuestión, directamente sobre la imagen de la cámara desplegada en tiempo real. Mediante esta tecnología es posible programar "secuencias tipo" de operación de cada equipo médico, las cuales son desplegadas gráficamente sobre la imagen del equipo. El sistema prototipo se desarrolló en lenguaje Java2 ME, y se ha evaluado sobre equipos Nokia. En trabajos futuros se pretende portar la aplicación para PDAs y evaluar su uso masivo.

Existe un gran potencial para esta aplicación dada la masificación y bajo costo de los dispositivos móviles, lo que permite vislumbrar un futuro cercano donde cada personal de los hospitales pueda visualizar y confirmar ciertos procedimientos en equipos médicos, simplemente "observando" a través de su celular.

Palabras clave— Realidad Aumentada, Procesamiento de Imágenes, Capacitación.

I. INTRODUCCIÓN

En los últimos cinco años se ha desarrollado el fenómeno de masificación mundial de la telefonía móvil, seguido por el creciente uso de los asistentes digitales personales (PDAs). Las cifras respecto de la cobertura mundial de estas tecnologías son extraordinarias. Por ejemplo, a nivel latinoamericano los países con mayor porcentaje de teléfonos móviles por persona, son Chile con un 74% y Argentina con un 68%. Sin embargo, si se analiza según grupos etarios, en la población joven y adulta los porcentajes son muy superiores. La tendencia al respecto la

presenta Europa, donde según estudios recientes un 95% de los jóvenes europeos posee un equipo celular.¹

El mercado de asistentes personales digitales (PDAs) presenta un crecimiento similar, del orden del 32% anual. Las ventas mundiales pronosticadas para el 2006 superan los 15 millones de unidades.²

Pero frente a todas estas cifras de cobertura y masificación de los equipos móviles, todo indica que posterior al fenómeno de la masificación de las tecnologías, debería irrumpir el fenómeno de la masificación de las aplicaciones. Hoy en día son pocas las aplicaciones móviles masivas (voz, mensajes de texto, mensajes multimedia, juegos, agendas, etc.). El acceso a Internet a través de equipos móviles presenta la mayor tendencia al alza, donde Japón lidera este crecimiento con un 92% de los usuarios de teléfonos móviles que utilizan sus equipos para navegar o revisar correos electrónicos³. La tendencia mundial al uso de equipos móviles debería ser similar. Recientemente se han cumplido 25 años desde el primer computador personal (PC), y son justamente algunos de los creadores del primer PC, entre los que destaca D. Bradley quienes vaticinan la creciente sustitución del PC por los dispositivos móviles.⁴

Los equipos móviles poseen cada vez mayores capacidades de procesamiento, aumento significativo en la capacidad de almacenamiento, e incorporan múltiples sistemas periféricos como canales de comunicación inalámbrica (Infrarrojo, Bluetooth, 802.11), cámaras, micrófonos, altavoz, entre otros. Junto a esto, las nuevas capacidades de telefonía en las PDAs, hacen prever un futuro cercano donde tanto las tecnologías de los teléfonos celulares como las tecnologías de las PDAs, confluyen en un dispositivo único y de uso masivo, que posee todas las características mencionadas y un grado de masificación global. En este contexto el "equipo móvil" de los próximos años se convertirá en el dispositivo más utilizado para el intercambio de información. Las capacidades de transducción de estos equipos no se limitarán a la digitalización de la voz, sino también a la identificación por radiofrecuencia (RFID), procesamiento de imágenes, lectores de códigos de barra, sistemas de reconocimiento de caracteres, sistemas de proyección de pantallas murales, etc.

Estas capacidades de los equipos móviles de nueva generación, permiten vislumbrar un campo prácticamente ilimitado de aplicaciones.

¹ <http://www.mundoenlinea.cl>

² Consultora Internacional Gartner

³ Estudio Ipsos Insight

⁴ <http://www.chicagotribune.com>

Todo lo anterior, concuerda absolutamente con los requerimientos de acceso a información de un campo muy especializado, como es la gestión y operación de equipos médicos. Las necesidades en esta área son múltiples, pero se puede enumerar algunas de las principales:

- Necesidad de información referente al modo de uso básico de los equipos.
- Necesidad de información referente a secuencias de operación y controles específicos de los equipos.
- Necesidad de información referente a marcas, modelos y proveedores de los diferentes equipos.
- Necesidad de información referente a la mantención y gestión de los equipos médicos.
- Se requiere minimizar los errores de digitación en el acceso a información
- Se requiere disminuir los tiempos de acceso a la información.

Las restricciones también son relevantes:

- En los campos clínicos (hospitales, centros médicos, etc.) tanto el personal técnico como el personal de la salud requieren un alto grado de movilidad, por lo tanto no pueden estar supeditados a portar diferentes equipos de lectura o computadores portátiles.
- La información generalmente se requiere *in situ*, y en tiempo real. No es factible postergar el acceso a la información, ni generar sistemas centralizados de acceso a la información. Tampoco es eficiente replicar documentos y manuales técnicos en las diferentes unidades de los recintos.

El uso de equipos móviles con cámaras incorporadas junto con el desarrollo de aplicaciones especiales permitiría responder a todas estas necesidades, cumpliendo además con las restricciones impuestas. El sistema propuesto consiste en identificar los diferentes equipos con un código visual único, mediante etiquetas adhesivas. Este código visual es capturado en tiempo real por la cámara del equipo móvil, y posteriormente analizado por el algoritmo de decodificación. Una vez obtenido el identificador del equipo, la aplicación puede desplegar en pantalla, diferentes tipos de información asociada al equipo. En primer lugar se puede desplegar información básica como la marca, modelo, fecha de adquisición, tiempo desde la última mantención, etc. Y en segundo lugar, es posible desplegar en la pantalla del equipo móvil la imagen en tiempo real entregada por la cámara y superponer textos, flechas o indicaciones visuales, asociadas directamente a los controles del equipo visualizado. Por ejemplo, la aplicación puede iniciar una secuencia de "repasso" de información para el usuario, donde se destaque cada uno de los controles y se indique su función, como se observa en la Fig. 1. Las aplicaciones que utilizan imágenes de video directo con leyendas u objetos superpuestos en base a referencias del "mundo real" se denominan aplicaciones de *realidad aumentada*. (Ver Apéndice B)

En este contexto, el objetivo de este trabajo, es determinar la factibilidad técnica de realizar aplicaciones de realidad aumentada en dispositivos móviles, que permitan obtener información en tiempo real, respecto de equipos u otros objetos. Esta información disponible en todo instante a través de los equipos móviles, permitiría facilitar muchos aspectos básicos de gestión y uso de los equipos.

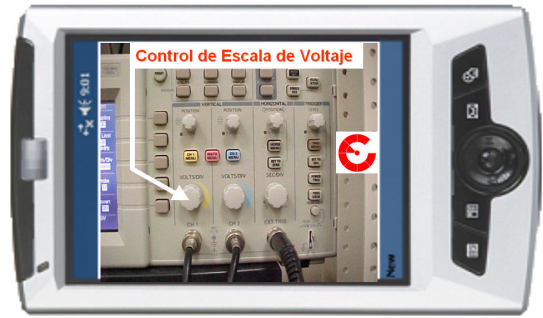


Fig. 1: Representación gráfica del sistema propuesto ejecutado en una PDA.

II. MATERIALES

Dada la gran diversidad de equipos móviles, es conveniente utilizar un entorno de programación flexible y de gran portabilidad. Es por ello que se utilizó JAVA como lenguaje base para los desarrollos. La mayoría de los equipos móviles poseen compatibilidad con máquinas virtuales de Java. Esto permite que un mismo programa desarrollado en Java pueda funcionar de manera similar en equipos con diferente hardware e incluso con diferente sistema operativo. Se utilizó una edición especial del lenguaje que se denomina Java 2 Micro Edition (J2ME). Para el acceso de las funciones de captura de imágenes se utilizó la librería MMAPI (Mobile Media API). Las rutinas de procesamiento de imágenes fueron desarrolladas íntegramente en Java y no se utilizaron librerías externas.

A pesar de lo reciente de estas tecnologías ya existen decenas de plataformas de desarrollo disponibles para la programación de dispositivos móviles. Específicamente para el desarrollo de este trabajo se utilizaron los siguientes entornos de desarrollo: Wireless Tool Kit 2.2 de SUN, y JBuilder Foundation de Borland, todos de libre distribución en Internet.

Los equipos móviles utilizados para las pruebas preliminares fueron el NOKIA 6620 y NOKIA 6670, que poseen una pantalla color de 176x208 píxeles de resolución, y hasta 65,536 colores, con Sistema Operativo Symbian 7.0. Además poseen una cámara CMOS de 3.5mm de distancia focal que permite capturar 15 cuadros por segundo, y hasta 1Mega píxeles de resolución.

Es importante destacar que existen restricciones relevantes en el desarrollo de software para equipos móviles. Específicamente en el entorno de desarrollo, Java J2ME (MIDP2.0/CLCD1.0), no existe la posibilidad de definir variables de tipo punto flotante ni operaciones trigonométricas. Todas estas operaciones se deben implementar con métodos alternativos utilizando sólo números enteros.

Sin embargo, la factibilidad de realizar procesamiento digital de imágenes en tiempo real sobre equipos móviles, ha sido comprobada en diferentes desarrollos previos de los autores. Además, en la literatura existen muchas publicaciones que respaldan las posibilidades reales de hacer procesamiento en dispositivos móviles. En particular destaca la publicación de J.Tiern [1] pues determina las capacidades y límites de procesamiento de imágenes de un equipo móvil estándar para este tipo de aplicaciones.

Es importante desatacar que cada día, los celulares y PDAs disponibles aumentan sus capacidades gráficas y de procesamiento, lo que augura cada vez mejores resultados de estas técnicas.

III. MÉTODOS

El sistema propuesto se puede analizar en base a los siguientes módulos de software:

A. El sistema de codificación

Existen múltiples sistemas de codificación en una o más dimensiones, como los conocidos códigos de barra. Sin embargo, muchos de estos estándares fueron desarrollados para sistemas de lectura láser y no son adecuados para decodificadores basados en cámaras. (Ver Apéndice A)

Para esta aplicación se desarrolló una familia de códigos visuales como se observa en la Fig.2.

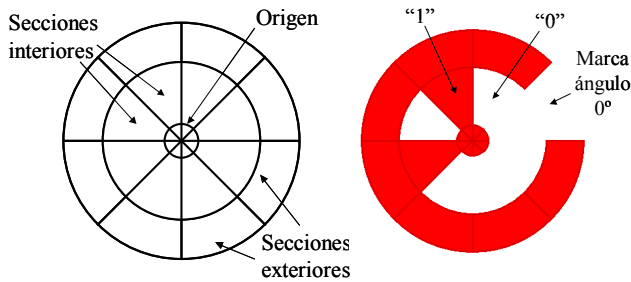


Fig. 2: a) Definiciones del Código Visual. b) Ejemplo de un Código Visual.

Estos códigos permiten una fácil detección mediante algoritmos de procesamiento de imágenes y fueron diseñados para proporcionar información respecto de un identificador numérico, como también información respecto de la distancia de captura. La identificador numérico está contenido en forma binaria, en base a la presencia o ausencia de color (rojo) en las diferentes secciones interiores del círculo. La distancia de captura o distancia entre el código visual y el equipo móvil, se determina en base al diámetro promedio del código circular de tamaño conocido.

El número de secciones interiores determina la capacidad numérica del código (2^n). Para ocho secciones como se observa en la Fig.2a, al igual que un número binario se puede codificar desde 0 hasta 255. Dependiendo del tipo y resolución de la cámara del equipo móvil, es posible lograr analizar códigos de mayor capacidad. Por ejemplo para 16 secciones interiores el código visual puede contener 65536 identificadores diferentes, lo cual permitiría a un hospital de gran envergadura asignar un código único a todos sus equipos médicos y otros equipos o materiales menores.

B. El algoritmo de decodificación

El algoritmo posee una serie de pasos secuenciales, descritos en la Tabla I, que permiten ubicar el código visual en la imagen capturada por la cámara, determinar las coordenadas de su origen, realizar la decodificación, y obtener la distancia de captura en base al diámetro promedio del código. Esta es la información base que permite el funcionamiento del siguiente módulo del sistema (la superposición de objetos virtuales).

La Tabla I, presenta los principales pasos del algoritmo y una breve descripción. En la Fig.3 se observan las diferentes imágenes de salida de cada uno de los pasos descritos en la Tabla I.

TABLA I

SECUENCIA DE OPERACIONES DEL ALGORITMO DE DECODIFICACIÓN.

Operación	Descripción	Imagen
Captura	Captura Imagen Original en Color	1
Filtro Por Color	Se potencia la detección del código visual "rojo" mediante la transformación $Gris=2*R-G-B$	2
Subsampling	Para aumentar la velocidad de procesamiento se disminuye la resolución	3
Binarización 1	En base a un umbral se binariza la imagen para determinar los potenciales códigos presentes	4
Histograma	Se realiza un histograma de potenciales centroides de círculos presentes en la imagen. Se asume que el mayor corresponde al código visual buscado	5
Zoom	Una vez ubicado el código visual, se realiza un <i>zoom</i> en dicha zona, para el posterior análisis del código	6
Binarización 2	Sobre la imagen ampliada se realiza una binarización y se determina el centroide del círculo.	7
Barrido	Se realiza un barrido mediante una recta en coordenadas polares, a partir del centroide. Cada recta representa el perfil de "corte" para un ángulo específico	8
Distancia 1	En cada barrido se determina la distancia hasta el último cruce de "1" a "0".	9
Prom Móvil 1	Se realiza un promedio móvil de las distancias hasta el último cruce y se determina así el mínimo, que corresponde a la "marca ángulo 0°"	10
Prom. Móvil 2	De manera similar se realiza el promedio móvil de la distancia hasta el primer cruce de "1" a "0". En este caso los peak corresponden a la presencia de secciones interiores del código	11
Detección	La imagen final indica gráficamente la detección de la "marca ángulo 0°" y de la sección interior en "1"	12

C. El procedimiento de superposición de información

Junto con el despliegue continuo de la imagen capturada por la cámara, el sistema obtiene a partir de la decodificación el identificador del equipo y la distancia y posición respecto del código visual. En base a esta información es posible superponer indicaciones (flechas y textos) que permiten al usuario obtener información respecto de cada uno de los controles del equipo. Para ello debe existir una base de información asociada a cada equipo, la cual fue estandarizada según el formato presentado en la Fig. 4, basado en el lenguaje XML.

La información asociada a cada identificador de equipo es guardada en un archivo de texto simple, cuyo nombre corresponde al identificador y la extensión ".xml". De esta manera para generar nuevos documentos sólo se requiere seguir las reglas básicas del lenguaje XML y utilizar un simple editor de texto.

El tag "<info>" enmarca toda la información general asociada al equipo. Dicha información es desplegada en la parte superior de la pantalla del software.

El tag "<panel>" se utiliza para indicar todos los controles (ctrl) que se desea incorporar al sistema de realidad aumentada. El formato inicial propuesto es muy básico y consiste en agregar tres atributos a cada control: el número "n" o identificador del control; y la posición x e y en centímetros respecto del origen del código visual. El código visual corresponde a la coordenada 0,0 y el factor de escalamiento para transformar las unidades de píxeles a

centímetros se obtiene del algoritmo de decodificación, mediante el cálculo del diámetro promedio del código, cuyo valor es conocido.

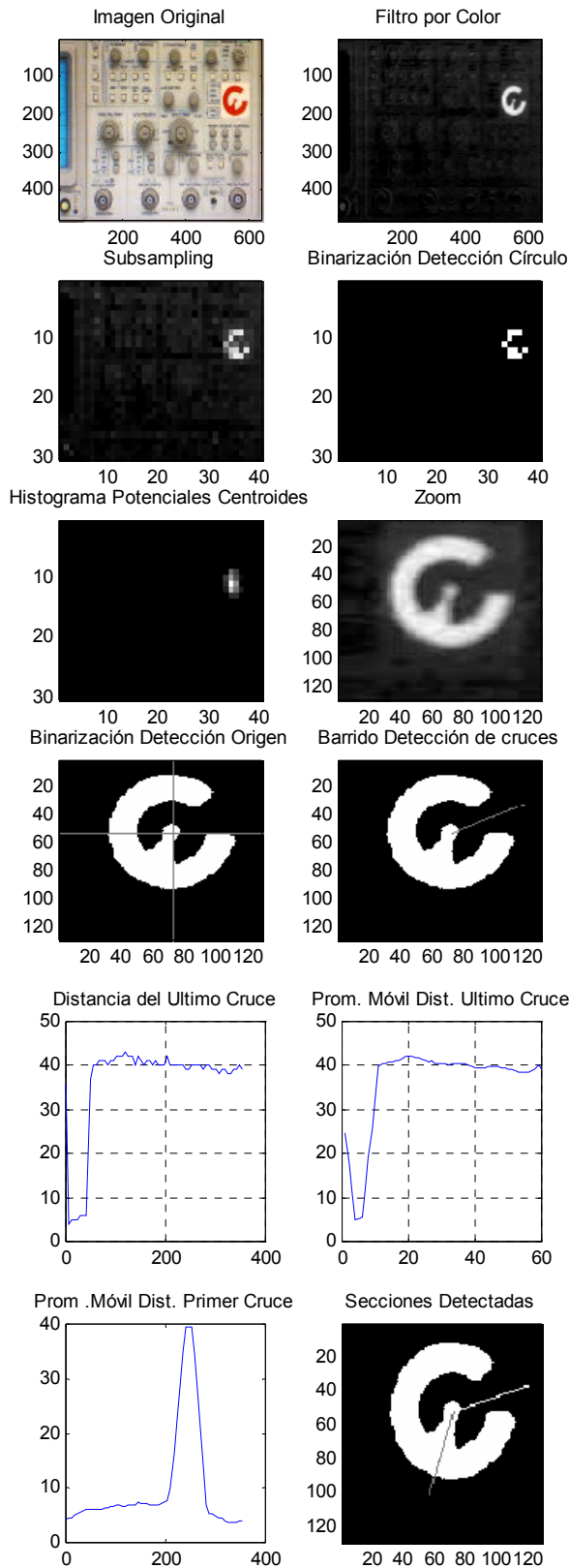


Fig. 3: Secuencia de imágenes resultantes del algoritmo de decodificación. Imagen 1 a la 12.

```

Nombre Archivo: 00000001.xml
<info>
  <marca>marca X</marca>
  <modelo>modelo X1</modelo>
</info>
<panel>
  <ctrl n=1 x=-30 y=+15>Encendido</ctrl>
  <ctrl n=2 x=-20 y=-10>Alarma</ctrl>
  <ctrl n=3 x=-18 y=-20>Volumen</ctrl>
</panel>

```

Fig. 4: Formato del documento con información asociada a cada equipo.

IV. RESULTADOS

El sistema desarrollado logra un correcto análisis del código visual y una adecuada superposición de marcas (flechas), para códigos (etiquetas) de tamaño 2cm^2 con color rojo sobre fondo blanco. La distancia necesaria para lograr correctas decodificaciones es de aproximadamente 60 cm entre el equipo móvil y el código visual. Todas las pruebas fueron ejecutadas sobre un teléfono Nokia 6670.

El tiempo de decodificación depende de la configuración de la cámara, sin embargo, se considera adecuada una resolución de 320×240 píxeles para las capturas. Considerando este formato de imágenes, el tiempo de decodificación es menor a 1 segundo, lo que permite una visualización continua de los equipos.

La Fig. 5 presenta una fotografía del sistema en ejecución, donde se despliegan diversos textos con información asociada al equipo decodificado.

V. DISCUSIÓN

Es importante tener presente que los equipos móviles presentan una serie de beneficios asociados precisamente a su portabilidad y autonomía, sin embargo, poseen también, varias restricciones que se pueden resumir en [2]: recursos de hardware y software limitados; vulnerabilidad un problema inherente a su portabilidad; conectividad inestable y variable ancho de banda; y energía finita. No obstante los resultados indican que con los celulares y PDA actuales es posible realizar aplicaciones de procesamiento digital de imágenes en tiempo real, por lo que se puede vislumbrar que en los próximos años, tanto los equipos como las aplicaciones de este tipo deberían aumentar aun más su rendimiento, precisión y velocidad.

VI. CONCLUSIONES

Respecto al objetivo del trabajo, se puede indicar que las principales ventajas del uso de equipos móviles como apoyo al acceso de información relativa a Equipos Médicos, radican en la simpleza y portabilidad del sistema, junto con su uso intuitivo y flexible. Existe un gran potencial para esta aplicación dada la masificación y bajo costo de los dispositivos móviles, lo que permite vislumbrar un futuro cercano donde tanto el personal técnico como el personal de la salud de los hospitales, pueda visualizar y confirmar ciertos procedimientos e información respecto de equipos médicos, simplemente “observando” a través de su celular.

VII. TRABAJOS FUTUROS

En trabajos futuros se pretende portar la aplicación para PDA's y evaluar su uso masivo.



Fig. 5: Imagen de la aplicación en ejecución.

APÉNDICES

A continuación se presentan antecedentes relevantes de dos temáticas fundamentales de este proyecto, como son los códigos visuales y la realidad aumentada.

A. Códigos Visuales

Como evolución natural de los conocidos códigos de barra, en las últimas décadas se han desarrollado numerosos estándares de códigos 2D (Fig.6). La mayoría de ellos han sido desarrollados para sistemas de lectura láser, por lo tanto no son adecuados para ser decodificados utilizando cámaras digitales (MaxiCode, DataMatrix, Aztec, PDF417 y otros). Sin embargo, existen otros códigos que se han desarrollado especialmente para sistemas de lectura basados en teléfonos móviles (Shotcode, ColorZip, eZcode, QRCode). A diferencia de éstos, el código visual propuesto en este trabajo, se caracteriza por permitir una correcta decodificación aun cuando el tamaño del código dentro de la imagen capturada sea muy pequeño. Esto facilita las aplicaciones de realidad aumentada, pues permite visualizar simultáneamente el objeto o escena real y el código visual. Por el contrario la mayoría de los códigos visuales mencionados, requieren un acercamiento mayor.

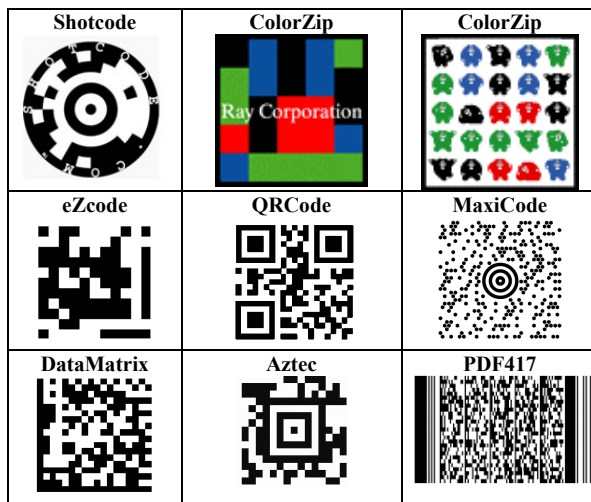


Fig. 6: Ejemplos de códigos visuales y códigos de barra 2D.

B. Realidad Aumentada

En los últimos 5 años, muchos de los investigadores que durante los años '90 realizaron aplicaciones en el campo de la realidad aumentada, han portado sus desarrollos a equipos móviles y celulares.

Cato and Billinghurst [3], utilizan target 2D para realizar tracking y para calibración de las cámaras en sistemas de visión con realidad aumentada. Henrysson, Billinghurst, Ollila [4] implementaron un sistema de colaboración "Face to Face" con Realidad Aumentada para equipos móviles, utilizando un target rectangular. Estos trabajos así como muchos otros similares se basan en el proyecto ARToolKit. Este proyecto emblemático en el ámbito de la Realidad Aumentada, fue inicialmente desarrollado por el Dr. Kato de Osaka University, y posteriormente fue apoyado por "Human Interface Technology Laboratory" (HITLab) de la Universidad de Washington, y HIT Lab NZ de la Universidad de Canterbury.

Grupos de investigación del INRIA, en Francia, han generado diversos trabajos de aplicación de la realidad aumentada. Hachet, Pouderoux, y Guitton [5] presentaron un manipulador de objetos 3D mediante el movimiento de un target frente a la cámara de una PDA.

Möhring, Lessing y Bimber [6] implementaron un sistema de realidad aumentada para visualizar objetos 3D a través de celulares. Este método fue estudiado anteriormente por Vallino & Kutulakos [7], pero utilizando cámaras convencionales.

Un conjunto muy interesante de aplicaciones de procesamiento de imágenes con celulares, fue desarrollada por Michael Rohs [8] del ETH / Suiza. Las diferentes aplicaciones han demostrado la factibilidad de realizar procesamiento de imágenes, en tiempo real y con gran precisión.

REFERENCIAS

- [1] J. Tierno, C. Campo, "Smart Camera Phones: Limits and Applications", *IEEE Pervasive Computing*, Abr-Jun 2005.
- [2] M. Satyanarayanan, "Mobile Information Access", *IEEE Personal Communications*, Vol. 3, No.1, Feb.1996.
- [3] H. Kato, M. Billinghurst, "Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System". In *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99)*. October, San Francisco, USA.
- [4] A. Henrysson, M. Billinghurst, M. Ollila: "Face to Face Collaborative AR on Mobile Phones". *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)* pp. 80-89
- [5] M. Hachet, J. Pouderoux, P. Guitton. "A Camera-Based Interface for Interaction with Mobile Handheld Computers", *Proceedings of I3D'05 - ACM SIGGRAPH 2005 Symposium on Interactive 3D Graphics and Games - 2005*
- [6] M. Möhring, C. Lessig, O. Bimber, "Optical Tracking and Video See-Through AR on Consumer Cell Phones". In *proceedings of Workshop on Virtual and Augmented Reality of the GI-Fachgruppe AR/VR*, pp. 193-204, 2004
- [7] J. Vallino, K. Kutulakos, "Augmented reality using affine object representations". *Fundamentals of Wearable Computers and Augmented Reality*, pp. 157 - 182, ISBN: 0-8058-2902-4, Lawrence Erlbaum (publisher), 2001
- [8] M. Rohs, P. Zweifel, "A Conceptual Framework for Camera Phone-based Interaction Techniques" *Pervasive Computing: Third International Conference, PERVASIVE 2005, Lecture Notes in Computer Science (LNCS)* No. 3468, Springer-Verlag, Munich, Germany, May 8-13, 2005