

# Detección automática de situaciones peligrosas para el telecuidado de personas mayores

José M. Cañas, Antonio Pineda, Pablo Barrera  
*Departamento de Ingeniería Telemática y Tecnología Electrónica*  
*Universidad Rey Juan Carlos*  
*28933 Móstoles*  
*jmplaza@gsync.escet.urjc.es*

## Resumen

*A la hora de desarrollar aplicaciones de telecuidado la visión proporciona un sensor incomparablemente bueno. Por un lado es poco intrusivo y además proporciona una gran cantidad de información a un precio bajo. Para su uso en telecuidado es necesario trabajar en tiempo real lo que implica una gestión ágil de las imágenes. Además, las imágenes sólo nos dan información bidimensional pero la mayor parte de la información útil es tridimensional, por lo que debemos ser capaces de extraerla de las imágenes. En el presente artículo presentamos una aplicación para el cuidado de personas mayores empleando únicamente información visual. La aplicación permite detectar problemas como pueden ser un desmayo, una caída al suelo o el acercamiento a una ventana. La monitorización automática de estas personas puede proporcionarles un mayor grado de autonomía.*

## 1. Introducción

El cuidado de personas mayores o enfermas implica una monitorización continua de su quehacer diario. En muchos casos son los propios familiares o los servicios sociales los encargados de cuidarles en los domicilios de estas personas o en residencias especializadas. Pero aún contando con el personal necesario para el cuidado es imposible observar a estos pacientes continuamente con el fin de detectar lo antes posible cualquier tipo de incidencias. El problema se agrava aún más en el caso de personas que viven solas, las cuales necesitan aún más este tipo de asistencia.

Los sistemas de teleasistencia tradicionales emplean terminales o dispositivos especiales que transportan los propios pacientes, en forma de pulsera o collar, y que en un momento de necesidad son accionados por éstos para enviar la llamada de emergencia a los servicios

necesarios. Sin embargo, estos sistemas tienen la limitación de que necesitan ser accionados por el usuario. En los casos más graves esto no será posible, con el consiguiente riesgo potencial que entraña. Por otro lado, el hecho de tener que cargar con un aparato de manera continua puede convertirse en un engorro para el usuario dada la naturaleza intrusiva del dispositivo.

Una alternativa para este tipo de servicios de telecuidado es la utilización de dispositivos de monitorización externos. Un conjunto de cámaras pueden encargarse de vigilar al paciente de manera automática. Cuando éste haga una señal o, más importante aún, cuando se observe algún comportamiento anómalo, como puede ser una caída al suelo, el sistema automáticamente puede enviar una señal de emergencia.

Este tipo de sistemas de monitorización automática basados en visión ha ido cobrando importancia en los últimos años. Por ejemplo en entornos sanitarios hay sistemas visuales para la detección de anomalías en radiografías, y en entornos biomédicos hay sistemas de localización 3D para evaluar desviaciones en la columna o perfeccionar los movimientos de atletas. En parte esto se debe al aumento de la capacidad de cómputo de los ordenadores y al descenso del precio de las cámaras. Cada día es más común encontrar multitud de cámaras a nuestro alrededor gracias a la economía de escala. Están presentes en los sistemas de vigilancia, en los móviles, en los ordenadores, en las consolas, etc.

Aun con tantas cámaras disponibles en la actualidad hay pocas aplicaciones totalmente autónomas basadas exclusivamente en visión. El motivo principal es que resulta muy costoso extraer información útil de las cámaras, lo que encarece sobremanera tanto el proceso de desarrollo de la aplicación como la fabricación de los dispositivos. Para que el sistema resulte útil en la

monitorización de personas enfermas debe ser capaz de funcionar de la forma más ágil posible con el fin de disparar la alarma con máxima celeridad. Además, tiene que trabajar de manera continua, durante todos los días. Conseguir este objetivo con recursos limitados impone restricciones al tipo de sistema a desarrollar y al tipo de tecnología que podemos emplear.

Para terminar, debemos considerar el tipo de información que resulta útil en un sistema de detección de situaciones peligrosas como el que aquí presentamos. En muchos casos uno de los principales problemas a la hora de identificar situaciones peligrosas consiste en que se emplea información visual de dos dimensiones. En el marco del análisis de imágenes puede resultar muy complejo encontrar o discernir entre una situación u otra. Un ejemplo puede ser si una persona se encuentra cerca o no de un fogón encendido. Al usar únicamente una imagen plana existe una ambigüedad a la hora de establecer la distancia por lo que podemos incurrir en errores fácilmente. Los requisitos de las aplicaciones para evitar estos errores hacen que incurramos en un alto coste computacional.

En este artículo presentamos una alternativa que se centra en la extracción de información tridimensional relevante de una escena. Una gran parte de la información que puede resultar útil del mundo que nos rodea es eminentemente tridimensional, como puede ser la posición relativa de un objeto frente a otro o el movimiento de una persona. Al usar información 3D resulta más sencillo describir las situaciones potencialmente peligrosas por lo que se simplifica en gran medida la aplicación, centrando la dificultad sobretodo en la extracción de información tridimensional.

Utilizando esta información tridimensional y un conjunto de reglas sencillas definidas en un mundo 3D hemos desarrollado una aplicación para el cuidado de personas mayores capaz de detectar en tiempo real situaciones potencialmente peligrosas, como puede ser un desvanecimiento, una caída al suelo o el acercamiento a una ventana. Todo ello se implementa con un conjunto de cámaras baratas, como puede ser unas *webcams*, y utilizando un PC de sobremesa corriente.

El sistema de extracción de información 3D que empleamos es novedoso y está diseñado para trabajar con unos recursos computacionales muy escasos y en entornos muy heterogéneos. Aunque ya existen aplicaciones que utilizan únicamente información

tridimensional usando imágenes [1], los métodos empleados para extraer dicha información suelen implicar un gran gasto computacional. En parte por esto, su aplicación no ha terminado de despegar más allá de ambientes muy controlados o con grandes recursos a su disposición, como pueden ser las cadenas de producción industrial o la creación de efectos visuales en la producción de películas o videojuegos.

## **2. Detección automática de situaciones peligrosas**

Cuando una persona mayor o enferma se encuentra sola puede ser útil tener un sistema que la monitorice con el fin de detectar posibles situaciones peligrosas lo antes posible. Un sistema así puede ayudar a aumentar su grado de autonomía, al no depender de que familiares o personal cualificado esté continuamente pendiente de él.

El sistema desarrollado está precisamente orientado a personas mayores que son susceptibles de padecer caídas o desmayos pero que no quieren vivir en una residencia de ancianos y prefieren seguir viviendo en sus casas, disfrutando de la mayor calidad de vida que eso supone. Un sistema como el que se presenta aquí permite reaccionar rápidamente en caso de esas situaciones peligrosas (caídas con pérdida de conocimiento, desfallecimientos, etc.) que se escapan a los sistemas actuales de teleasistencia mediante pulsador.

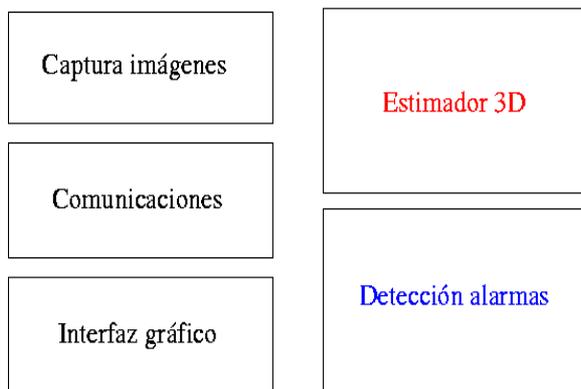
Para que una aplicación así sea realmente útil debe funcionar autónomamente. Es decir, debe funcionar sin necesidad de que haya otra persona supervisando permanentemente, ya sea en el mismo lugar que el usuario o en una central del control remota. El sistema presentado aquí es completamente autónomo. Al ser automático, aumenta la comodidad de los familiares o los cuidadores, que no tienen que estar permanentemente pendientes. De hecho, el personal externo sólo será necesario para atender a las emergencias detectadas por la aplicación, no de continuo.

Adicionalmente, el empleo de cámaras en el hogar plantea un dilema ético importante en lo relacionado con la intimidad. Con un sistema automático como el diseñado se mantiene la intimidad de la persona cuidada, pues nadie le observa. No hay nadie vigilándole “al otro lado”. Las imágenes recogidas por

el sistema no se graban en ningún soporte y no salen del sistema salvo que lo justifique una situación de emergencia en la que la vida de la persona corra peligro. Se puede ver este sistema como un servicio automático de alarmas, similar a los sensores volumétricos que ya hay en muchos hogares y que efectivamente miden nuestra presencia en distintas estancias de una casa.

No obstante, al ser las cámaras unos sensores tan potentes, y existir la posibilidad de enviar las imágenes remotamente a los cuidadores externos, habría que incorporar al sistema mecanismos de seguridad que refuercen la protección frente a intrusiones. Si la señal de alarma se reduce a una llamada de teléfono o un mensaje SMS no habría posibilidad de invasión de intimidad.

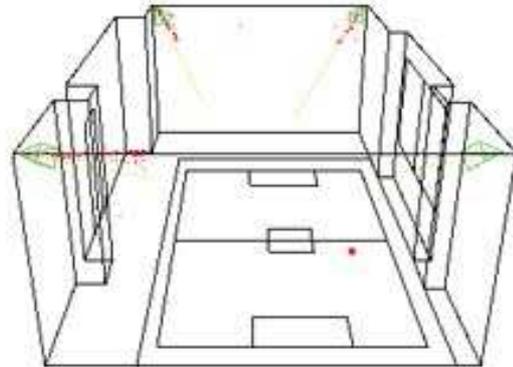
El sistema básico que hemos desarrollado está compuesto por varios módulos interconectados entre sí. Podemos ver en la Figura 1 un esquema con los módulos que componen nuestra aplicación.



*Figura 1: Diagrama de bloques de nuestra aplicación de telecuidado*

En primer lugar tenemos un sistema de captura de imágenes a través de un conjunto de cámaras, tanto móviles como estáticas, encargadas de observar al usuario continuamente. La colocación de las cámaras para una aplicación típica puede verse en la Figura 2. En ella se observan cuatro cámaras colocadas en cada

una de las esquinas de una habitación lo que nos permite ver cualquier punto de la misma. Las cámaras están conectadas, directamente o a través de la red, a un PC de escritorio convencional.



*Figura 2: Prototipo con 4 cámaras en los rincones superiores*

En el PC tenemos el módulo encargado de procesar las imágenes de manera independiente y combinar los resultados para obtener la información tridimensional.

Una vez localizada la persona dentro de la escena podemos usar esta información tridimensional para determinar si existe una situación de peligro. Podemos dividir la situaciones de peligro observables en dos grandes grupos. Por un lado tenemos aquellas relacionadas con lugares potencialmente peligrosos, como puede ser una terraza o una puerta para una persona con alzheimer. Por otro lado la posición en la que se encuentra la persona, si está encorvada, agachada o incluso se ha caído al suelo y no puede levantarse. Gracias a que tenemos información tridimensional de la posición de la persona podemos determinar si nos encontramos en alguna de estas situaciones que hemos definido como potencialmente peligrosas.

Una vez detectado que la persona continúa en una zona de peligro o, por ejemplo, no es capaz de levantarse, el sistema dará la voz de alarma automáticamente.



*Figura 3: El sistema ha detectado automáticamente la caída de la persona*

Al emplear un PC para procesar la información, resulta muy sencillo enviar el aviso por Internet o incluso, empleando telefonía móvil, llamar directamente al servicio de urgencia acompañándolo con una imagen o secuencia de vídeo para que un operador humano pueda evaluar la gravedad de la situación y así seleccionar los medios humanos necesarios para intervenir en esa situación.

### 3. Tecnología de seguimiento visual 3D

Para obtener información realmente tridimensional se emplean varias cámaras al mismo tiempo dirigidas hacia el usuario. Emplear más de una cámara permite realizar un seguimiento más fluido y, además, facilita la extracción de información tridimensional, con lo que se reduce el coste computacional. El número de cámaras necesarias dependerá del recinto que se desee cubrir. Será necesario que todas las zonas a seguir, por lo menos aquellas importantes, caigan dentro del campo visual de al menos dos cámaras. En caso contrario, no será posible obtener información precisa sobre el usuario, aunque todavía sería posible sacar

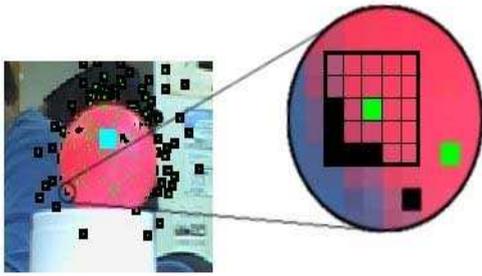
información indirecta, como puede ser el tiempo que ha pasado desde que se vio al usuario por última vez.

El procesamiento de las imágenes que realiza el sistema parte de la idea de que resulta más sencillo realizar hipótesis de la realidad y, posteriormente, comprobar cuales pueden cumplirse que intentar realizar una reconstrucción partiendo de las imágenes. En nuestro caso empleamos un filtro de partículas para realizar la labor de creación y búsqueda de posibles hipótesis [2,4,5]. Cuando se obtienen las hipótesis de la posición del objeto o de la persona que queremos buscar el sistema se encarga de capturar las imágenes y comprobar si la información captada por las cámaras coincide con las hipótesis.



*Figura 4: Ejemplo de seguimiento 3D con 4 cámaras*

En la figura 4 podemos ver las cuatro imágenes captadas por las cámaras. Superpuestas hay una serie de puntos rojos que representan a cada una de las hipótesis que realizamos sobre la posición de persona en la escena. Las hipótesis son un conjunto de posiciones reales en 3D y lo que se muestra en las imágenes son la proyección de dichos puntos sobre cada una de las cámaras. En caso de que el objeto esté en la misma posición que una de las hipótesis, ésta proyectará correctamente en cada una de las cámaras.



*Figura 5: Modelo de observación basada en color*

Para comprobar si la proyección cae o no dentro de una persona empleamos un sencillo filtro de color y otro de movimiento en el entorno del punto proyectado. Contando el número de píxeles que pasan el filtro tendremos una medida de la verosimilitud, aunque muy simplista, de que el objeto se encuentre en esa posición. En la figura 5 podemos ver un ejemplo de múltiples hipótesis sobre una pelota de color rosa. Aunque los filtros son relativamente sencillos, su combinación y, sobre todo, su persistencia en el tiempo resulta muy discriminante. En los casos en que el color no pueda emplearse como parte del discriminante se puede utilizar un esquema de aprendizaje de color o forma basado en el movimiento de nuevos objetos. Tan pronto como aparezca un objeto en movimiento en el escena el sistema aprenderá sus colores lo que permitirá configurar el filtro de color para seguirlo de manera adecuada.

Al emplear la combinación de color y movimiento podemos aunar las ventajas de cada uno de estos procesamientos, complementándose entre sí. Por un lado, el movimiento nos seguirá a las personas de manera ágil, dado que no suele haber muchos otros objetos en movimiento a su alrededor. En una casa normal los objetos están estáticos y únicamente son de interés para nuestra aplicación aquellos que se mueven. Un detector de movimiento discriminará de manera sencilla mucha de la información que recibimos a través de las cámaras. Su limitación viene cuando la persona está quieta, lo que la convierte en invisible a este tipo de observación. La información de color, por su parte, nos permite discriminar de manera fácil qué objetos nos resultan más interesantes usando una característica visual muy sencilla. La combinación de estos dos tipos de observaciones tiene la ventaja de

aprender el color cuando algo se mueve y de no perderlo cuando está quieto.

Gracias a la sencillez de estos filtros podemos comprobar decenas de miles de hipótesis por segundo en un PC de sobremesa cualquiera lo que nos da una idea de la vivacidad del sistema. En términos netos, esta sencillez mantiene el bajo el coste computacional del algoritmo de seguimiento

#### 4. Conclusiones y Líneas Futuras

Hemos presentado un sistema que detecta automáticamente situaciones peligrosas para el telecuidado de personas mayores y dispara la alarma cuando éstas se presentan. En particular, detecta autónomamente desmayos y caídas. A la señal de alarma se le puede asociar varias acciones como el envío de un SMS o de las imágenes al teléfono móvil de un familiar, la llamada a un médico, etc.

Nuestro sistema consta de varias cámaras de videoconferencia, un ordenador, y el software de procesamiento visual. El hardware es convencional, lo que permite su rápida sustitución en caso de deterioro, y mantiene bajos los costes económicos del sistema. El software se basa en avanzadas técnicas de seguimiento visual en tres dimensiones, los filtros de partículas.

Hemos construido un prototipo experimental y lo hemos probado detectando caídas en una habitación y disparando una alarma acústica cuando éstas se producían. El seguimiento en tres dimensiones conseguido es muy vivaz, lo que permite una detección inmediata de la situación peligrosa y con ello, una respuesta más ágil a la emergencia.

Frente a los dispositivos colgantes que el propio anciano pulsa cuando necesita ayuda, nuestro sistema funciona incluso cuando la persona a monitorizar se desmaya y pierde la consciencia. Además, el usuario no tiene que hacer nada especial, ni llevar ningún dispositivo colgado.

Más allá de las aplicaciones de televigilancia, que simplemente trasladan las imágenes a un operador humano remoto, nuestro sistema detecta automáticamente la emergencia. Esto permite mantener la privacidad de las personas telecuidadas y permite realizar una monitorización continua, todos los días del año, pues el sistema "no se cansa".

Actualmente estamos trabajando en aumentar la funcionalidad del prototipo para el seguimiento

simultáneo de varios objetos, y extendiendo la definición de situaciones peligrosas.

## Referencias

- [1] Vicon Peak <http://www.vicon.com>
- [2] Barrera P., Cañas J. y Matellán V., “Visual object tracking in 3D with color based particle filter” , Int. Journal of Information Technology, Vol 2, Num 1, pp 61—65, 2005.
- [3] Fritsch J., Kleinhagenbrock M., Lang S., Fink G. y Sagerer G. “Audiovisual person tracking with a mobile robot”, Proceedings of Int. Conf. on Intelligent Autonomous Systems, pp 898—906, 2004.
- [4] Pérez P., Vermaak J. y Blake A. “Data fusion for visual tracking with particles”, Proceedings of IEEE, Vol 92, Num 3, pp 495—513, 2004.
- [5] Pupilli M. y Calway A. “Real-Time Camera tracking using a particle filter”, Proceedings of British Machine Vision Conference, pp 519—528, 2005.
- [6] Zotkin D., Duraiswami R. y Davis L. “Multimodal 3D tracking and event detection via the particle filter”, IEEE Workshop on detection and recognition of events in video, pp 20-27, 2001.