

DetECCIÓN Y SEGUIMIENTO DE OBJETOS EN VÍDEOS DE ACTIVIDADES DE VIDA DIARIA PARA REHABILITACIÓN DE PACIENTES CON DAÑO CEREBRAL ADQUIRIDO

M. Luna Serrano^{1,2}, P. Sánchez-González^{1,2}, E. Bonilla Carrión^{1,2}, A. Fernández Pérez^{1,2}, J.M. Martínez Moreno^{1,2}, Marc Morell Vilaseca³, C. Cáceres Taladriz^{1,2}, J.M. Tormos Muñoz³, E. J. Gómez Aguilera^{1,2,1}

¹ Grupo de Bioingeniería y Telemedicina (GBT), E.T.S.I. de Telecomunicación, Universidad Politécnica de Madrid, Madrid, España, {mluna, psanchez, ebonilla, afernandez, jmartinez, ccaceres, egomez}@gbt.tfo.upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ Instituto Guttmann – Hospital de Neurorehabilitación, Badalona, España, jmtormos@guttmann.com

Resumen

Las técnicas de rehabilitación permiten la recuperación y mejora de las funciones dañadas o deterioradas y ayuda al paciente con DCA a adaptarse a su nueva situación. El avance tecnológico que se ha producido en las últimas décadas, ha impulsado la investigación en el diseño y desarrollo de nuevos modelos de rehabilitación. La tecnología de video interactivo se convierte en un elemento de apoyo en estos nuevos modelos rehabilitadores. Se hace necesario desarrollar nuevos algoritmos de segmentación y seguimiento que permitan dotar de información adicional a los vídeos. En este trabajo se han implementado y evaluado dos métodos que permiten realizar la detección y el seguimiento de objetos de interés.

1. Introducción

El daño cerebral adquirido (DCA) hace referencia a cualquier lesión de las estructuras cerebrales como consecuencia de una enfermedad o accidente [1]. La lesión más frecuente es el traumatismo craneoencefálico (TCE) [2], cuya tasa de incidencia se sitúa alrededor de los 200 nuevos casos por cada 100.000 habitantes y año [3].

Las técnicas de neurorehabilitación permiten la recuperación de algunas de las funciones perdidas, la mejora de capacidades conservadas y la ayuda al paciente para adaptarse a sus limitaciones. De esta forma se logra mejorar la autonomía y la independencia de la persona afectada. Los avances tecnológicos pueden permitir la mejora, la personalización y el aumento de la eficacia de los procesos de rehabilitación actuales. En el proyecto REHABILITA [4], se están investigando nuevas formas y metodologías que supongan un cambio de paradigma en los procesos de rehabilitación de pacientes con DCA. El diseño, creación y aplicación de nuevos elementos rehabilitadores basados en tecnologías de video interactivo (VI) es uno de los aspectos que se están abordando en esta línea de investigación. Se entiende como VI cualquier aplicación de vídeo en el que la secuenciación y la visualización de la información se determina por la respuesta del usuario [5]. Así, el VI es una secuencia de fotogramas que se relacionan con otros archivos en función de la secuencia de acciones del

usuario y de un conjunto de descripciones que permiten gestionarlo de forma inteligente [6].

Actualmente, las características del VI permite contemplarlo como una tecnología ideal para el desarrollo de actividades de vida diaria (AVD) que cumplan con los pilares sobre los que se basa el nuevo paradigma de rehabilitación: personalización, ubicuidad y monitorización [4].

Uno de los puntos importantes en la creación de VI es la definición de las zonas calientes, es decir, aquellos segmentos de los fotogramas que se van a dotar de interactividad. Para ello es necesario el diseño e implementación de algoritmos de procesamiento de vídeo que permitan detectar y seguir objetos a lo largo de una secuencia de fotogramas [7] [8].

En este artículo se describen dos métodos que permiten la detección y seguimiento de objetos en vídeo. El objetivo final es la creación de entornos virtuales interactivos basados en vídeo interactivo que permitan una mejora en la rehabilitación y la incorporación temprana a la vida diaria de personas con DCA.

2. Material y métodos

En el presente trabajo de investigación se han desarrollado dos métodos de detección y seguimiento, uno semiautomático y otro automático. Estos métodos detectan y siguen los objetos que o bien son seleccionados por el usuario (método semiautomático) o se localizan sin la necesidad de la intervención del usuario (método automático).

El método semiautomático es necesario en entornos donde sea complejo realizar una descripción unívoca de un objeto y por tanto no sea posible obtener un patrón del mismo. El método automático se utiliza cuando es posible obtener un patrón unívoco que defina el objeto. De tal forma, cada uno de los objetos que son identificados automáticamente se encuentran almacenados en una base de patrones. El método de detección y seguimiento utiliza las características almacenadas en la base de patrones para identificar el objeto de interés en una determinada escena.

Los métodos propuestos se han desarrollado sobre C++ utilizando la librería OpenCV. Las pruebas de evaluación se han llevado a cabo en un Intel Core2 Duo a 3GHz, 4 GB de RAM y Windows 7 Professional 32 bit.

2.1. Método semiautomático basado en el ratio del histograma.

Este método se basa en extraer información del histograma de la imagen para llevar a cabo la detección y seguimiento de objetos en vídeo. En la Figura 1 se puede observar un diagrama de bloques de este método. El primer paso es el cálculo del ratio del histograma dividiendo el histograma en regiones del mismo tamaño que la región que caracterice el objeto seleccionado y seleccionando únicamente la región que contenga al objeto de interés. Este parámetro permite ignorar principalmente los colores del objeto que están presentes en el resto de la escena y se escala el histograma resultante.

En la siguiente etapa se aplica el algoritmo CAMSHIFT (*Continuously Adaptive Mean Shift*) [6], que utiliza la información del ratio estimado para calcular la región más probable del nuevo fotograma en la que se puede encontrar el objeto de interés. Finalmente, se resalta sobre la imagen la zona seleccionada a lo largo del intervalo. Las opciones de marcado varían en función de la aplicación del método. Así, se puede resaltar la máscara exacta mediante contornos o zonas de transparencia, o remarcarlo con un rectángulo que cubra el área de interés.

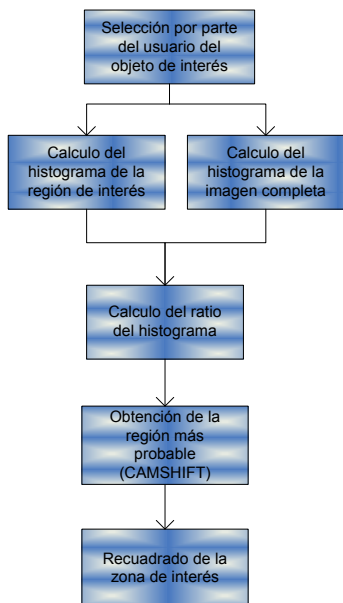


Figura 1 Diagrama de bloques del método semiautomático.

2.2. Método automático basado en las diferencias entre las componentes RGB.

El objetivo de este método es detectar y seguir de forma automática objetos que tengan unas determinadas características de color. En la Figura 2, se muestra un diagrama de bloques del método automático:

De cada uno de los fotogramas se extrae la componente de color del objeto que se desea segmentar. A continuación, se obtiene la imagen en escala de gris y por

último se halla la diferencia entre la componente seleccionada y la gris. Esto nos permite separar y aislar los objetos que contengan información relativa a la componente de color seleccionada.

Posteriormente, se elimina el ruido existente en la imagen mediante un filtro de media y se umbraliza la imagen para obtener una imagen binaria. En este punto, se emplean operaciones morfológicas para fusionar las zonas de menor tamaño, etiquetándose las regiones de interés, con el fin de resaltarlas en la imagen final.

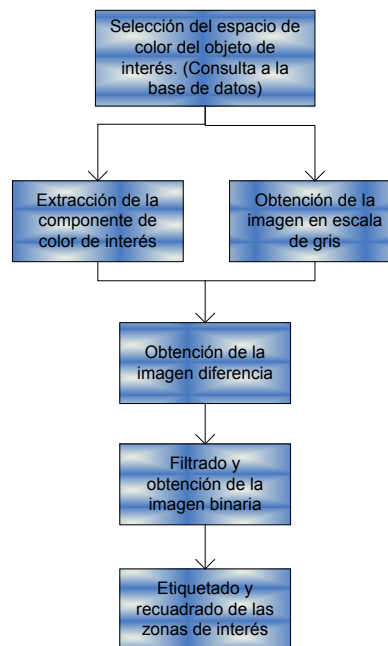


Figura 2 Diagrama de bloques del método automático

2.3. Metodología de evaluación

La metodología de evaluación de los algoritmos de procesamiento parte de una segmentación manual del objeto sobre todos los fotogramas, empleada como "gold estándar". Para la validación de ambos métodos se emplean los mismos vídeos.

Los parámetros que se van a utilizar para evaluar estos resultados están descritos en [9]: tiempo de ejecución, sensibilidad, especificidad, valor predictivo positivo (VPP) y valor predictivo negativo (VPN).

3. Resultados

Tal y como se ha comentado en el apartado anterior, la validación de ambos métodos se ha llevado a cabo con los mismos vídeos para que la comparativa sea de interés. Los resultados se particularizan para un vídeo perteneciente a la actividad de vida diaria de vestirse que tiene 101 fotogramas y una duración de unos 4 segundos.

3.1. Resultados obtenidos con el método de segmentación semiautomático basado en el ratio del histograma.

Los resultados cualitativos del método se observan en la Figura 3.

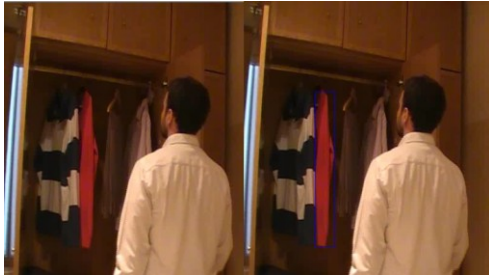


Figura 3 Resultado de aplicar el método basado en el ratio del histograma.

El tiempo de ejecución de este método es de 91,98 segundos. La fluctuación de la sensibilidad es grande y su valor medio es del 55%. Así, el método segmenta correctamente más del 50% de la región de interés que se desea segmentar. La especificidad es del 94,6%. Los resultados hacen ver que este método descarta correctamente los píxeles que no corresponden a la zona de interés de una forma muy fiable. El VPP medio es del 28%. El VPN presenta una alta variabilidad y un valor medio situado en torno al 98%, con lo que el método descarta correctamente los píxeles que no corresponden al objeto de interés.

En la Figura 4 se observa la evolución de los valores de cada uno de los parámetros que se han utilizado para realizar la evaluación.

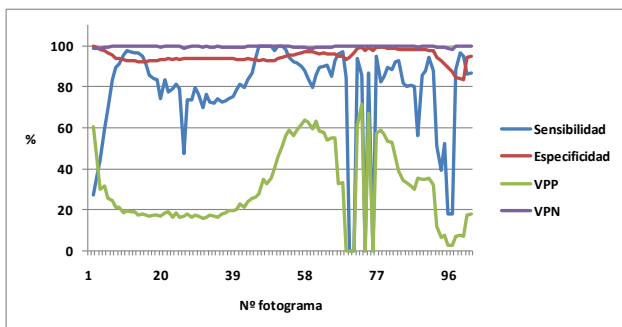


Figura 4 Evolución de los parámetros de evaluación al aplicar el método semiautomático basado en el ratio del histograma.

3.2. Resultados obtenidos con el método de segmentación automático.

El tiempo de ejecución de este método es de 1.26 segundos. La sensibilidad tiene un valor medio del 68,5%, es decir, se segmenta correctamente más del 68% de la región de interés. La especificidad presenta unos valores de especificidad muy altos, teniendo un valor medio de 99.5%. El VPP tiene una alta fluctuación, aunque su valor medio es del 85.9%. El porcentaje de píxeles descartados correctamente es muy alto, siendo el valor medio VPN del 98.3%.

En la Figura 5 se pueden observar las variaciones de estos parámetros en función del fotograma.

Como se puede observar en la Figura 6, la máscara o imagen binaria que se obtiene cubre prácticamente la totalidad del área del objeto de interés, que en este caso es un objeto de color rojo.

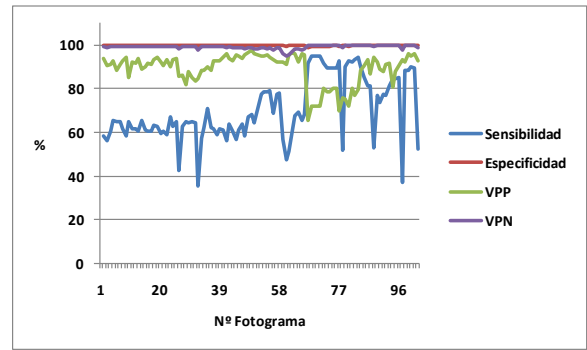


Figura 5 Evolución de los parámetros de evaluación al aplicar el método automático.



Figura 6 Resultado de aplicar el método automático.

4. Discusión

El avance tecnológico durante estas últimas décadas ha impulsado la investigación en el campo de la mejora de los modelos de rehabilitación. El vídeo interactivo aparece como una tecnología cuyas características la hacen propicia para su utilización dentro del nuevo paradigma de la rehabilitación. Esto hace necesario desarrollar nuevos algoritmos de procesamiento que permitan dotar de información a los vídeos. En este trabajo se han implementado y evaluado dos métodos que permiten realizar la detección y el seguimiento de objetos de interés.

Los métodos permiten detectar y seguir objetos a lo largo de un vídeo. Esto se empleará para la creación de regiones a las que se les dotará con propiedades interactivas que permitan la interacción del usuario, de forma que éste pueda intervenir en la reproducción del vídeo. Como resultado de asociar interacciones y regiones de los fotogramas se crearán diferentes flujos de acción en función de lo que ha sido diseñado en el árbol del VI.

El método semiautomático necesita de la interacción del usuario para seleccionar el objeto de interés. El método automático no necesita de esta interacción y su función es la de resaltar objetos de interés y de relativa importancia para el usuario.

Ambos métodos son necesarios en el contexto de aplicación actual, ya que es necesario detectar tanto objetos que se puedan definir de forma unívoca como otros objetos que presentan algo más de complejidad en su definición. Así, ambos métodos se usan al evaluar la capacidad de elección del paciente: de tal forma, si el objeto seleccionado es correcto se le mostrará un nuevo fotograma al paciente.

En la Tabla 1 se recogen los datos promedio obtenidos como resultado de aplicar cada uno de los dos métodos implementados.

	Método semiautomático	Método automático
Sensibilidad	55,5 %.	68,5 %
Especificidad	94,6 %	99,6 %
VPP	28,2 %	85,9 %
VPN	98,4 %	98,3 %
Tiempo	91,98 s	1,26 s

Tabla 1 Valores promedio de los parámetros de evaluación de los métodos implementados.

En los dos métodos el valor de especificidad y el VPN son muy altos. Los resultados, por tanto, hacen ver que estos métodos descartan correctamente los píxeles que no pertenecen a la estructura de interés.

Tanto la sensibilidad como el valor predictivo positivo (VPP) es mayor en el método automático, ya que la cantidad de píxeles segmentados que pertenecen realmente al objeto de interés es muy alta. Aunque estos valores son algo menores para el método semiautomático, se puede afirmar que este método realiza una correcta detección y segmentación del objeto de interés: el cálculo del ratio del histograma ajusta con bastante precisión los parámetros que se necesita buscar en los sucesivos fotogramas.

Como se puede deducir de los resultados obtenidos ambos métodos ofrecen buenos resultados al detectar y seguir objetos. Además, ambos métodos descartan correctamente los píxeles que no corresponden al objeto de interés, obteniéndose valores próximos al 100% en especificidad y VPN. Se obtienen también buenos valores de sensibilidad y VPP, aunque en este caso es el método automático el que obtiene mejores resultados se puede concluir que en ambos casos se selecciona correctamente la región que contiene el objeto de interés.

En cuanto a los tiempos de ejecución, se puede observar que el método que tiene un tiempo más alto es el método basado en el ratio del histograma. A pesar de ello, ambos métodos están dentro de los límites temporales del contexto de aplicación, ajustándose correctamente a las necesidades actuales.

Actualmente se están desarrollando nuevos métodos tanto automáticos como semiautomáticos con la finalidad de poder detectar objetos usando otros parámetros de las imágenes. Se está también analizando la viabilidad de utilizar dichos algoritmos en entornos de gestión inteligente de videos terapéuticos basados en sistemas de recuperación de contenido.

5. Conclusiones

El objetivo de la rehabilitación es la recuperación y mejora de las funciones dañadas, facilitando la adaptación a los pacientes que han sufrido algún tipo de DCA. Los avances tecnológicos han impulsado nuevos campos de investigación con el objetivo de mejorar y personalizar estos procesos de rehabilitación. Los videos interactivos con fines terapéuticos son uno de los avances tecnológicos que permiten una mayor personalización en este tipo de terapias.

En este trabajo se han implementado y evaluado dos métodos que detectan y siguen objetos de interés a lo largo de una secuencia de fotogramas. Estos métodos permiten dotar al video de interactividad e información, al posibilitar la definición de regiones a las que se les dota de propiedades interactivas. Los resultados obtenidos muestran su viabilidad de aplicación en el contexto de la investigación actual: facilitar la creación de entornos virtuales interactivos basados en VI con los que se alcance una rehabilitación personalizada, ubicua y basada en la evidencia.

Agradecimientos

Este trabajo de investigación ha sido parcialmente financiado por el proyecto CENIT-E “REHABILITA” CEN-20091043. Los autores expresan su agradecimiento a todos los socios del consorcio REHABILITA.

Referencias

- [1] Bernabeu Guitart M, Roig Rovira T. Unidad de daño cerebral en el Instituto Guttmann. Hospital de Neurorehabilitación. *Revista MinusVal*, 2002, pp. 6-8.
- [2] Instituto Guttmann: Hospital de Neurorehabilitación. Datos propios sobre la tasa de incidencia de daño cerebral adquirido. 2008.
- [3] Roig T. Rehabilitación Cognitiva en personas con daño cerebral adquirido. *En Sobre Ruedas*, N 76, 2010, pp. 13 – 14.
- [4] Tormos Muñoz JM, Royo Sánchez C, Gómez Aguilera EJ. REHABILITA: Tecnologías Disruptivas para la Rehabilitación del Futuro. *XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2010)*, Madrid, 2010 (ISBN: 9788480581).
- [5] Stephan S, Riempp R. The cognitive benefits of interactive videos: learning to tie nautical knots. *Science Direct*, 3(14), 2004, pp. 293 - 305
- [6] Bosco J. An analysis of evaluations of interactive video. *International Conference on Computers and Education*, 1990.
- [7] Jain AK, Duin RP, Mao J. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence* 22, nº 1, 2000, pp. 4–37.
- [8] Bradski GR. Computer vision face tracking as a component of a perceptual user interface. *In Workshop on Applications of Computer Vision*, Princeton, NJ, Oct. 1998, pp 214–219,
- [9] Pérez Rodríguez R, Solana Sánchez J, Cano González AM, Sánchez-González P, Gómez Aguilera EJ. Segmentación y seguimiento de estructuras en imágenes laparoscópicas. *Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB2008)*, 2008, pp.108-111.
- [10] Qian H et al. Chapter 3: Segmentation and Tracking. *Intelligent Surveillance Systems. Intelligent Systems, Control and Automation: Science and Engineering* 51, DOI 10.1007/978-94-007-1137-2_3. Springer Science Business Media B.V. 2011.
- [11] Raja Y, McKenna SJ, Gong S. Segmentation and Tracking Using Colour Mixture Models. *Lecture Notes in Computer Science*, 1997. vol. 1351/1997, pp. 607 – 614.