

Medidas de Evaluación para Métodos de Seguimiento de Objetos en Video

Ximena Pisani, Juliana Gambini

Instituto de Ciencias, Universidad Nacional de General Sarmiento
Los Polvorines, Pcia. de Buenos Aires, Argentina
xpisani@ungs.edu.ar mgambini@ungs.edu.ar

Resumen La evaluación de la eficiencia de un algoritmo de seguimiento de objetos en video es un desafío en el área de procesamiento de imágenes y adquiere mayor importancia cuando se trata de interpretación automática de secuencias de imágenes médicas, donde la precisión de los resultados es fundamental. Por otro lado, un sistema de medición de la performance de un método permite comparar diferentes algoritmos. En este trabajo de investigación se propone investigar diferentes medidas de evaluación del comportamiento de algoritmos de tracking que soporten oclusión. Comenzamos con la evaluación de un método rápido de seguimiento basado en conjuntos de nivel pero que no resuelve ecuaciones diferenciales sino que ajusta el contorno del objeto de interés en cada cuadro utilizando intercambio de pixels. La métrica utilizada está basada en el color de los pixels de la curva que forma el contorno.

Palabras Clave: Seguimiento de Objetos en video, Contornos Activos, Medidas de Evaluación del Comportamiento

1. Contexto

La línea de investigación presentada en este trabajo está inserta en el proyecto "FONDO SEMILLA 2010", acreditado y coordinado por

la Universidad Nacional de General Sarmiento. Dirigido por la Dra. Juliana Gambini.

2. Introducción

El problema de seguimiento de objetos en secuencias de imágenes y video tiene múltiples aplicaciones en el estudio del medio ambiente, estudio del comportamiento de animales, robótica, seguridad y diagnóstico médico, entre otras. Diversos autores han desarrollado algoritmos automáticos de seguimiento de objetos en video [11, 2, 1], incluso en presencia de oclusión [6, 3]. También existen algoritmos de tracking basados en conjuntos de nivel, como [21], [9], [15], [4], [14]. Todos ellos resuelven en cada paso ecuaciones diferenciales en derivadas parciales y resultan muy robustos, además permiten cambios en la topología pero tienen serias limitaciones en aplicaciones que requieren tiempo real. En los artículos [19, 20], los autores proponen un método de seguimiento de objetos en video, que utiliza la teoría de conjuntos de nivel pero en el cual no es necesario resolver ecuaciones diferenciales en derivadas parciales, sino que la evolución se realiza por medio de intercambio de pixels. En el artículo [20] los autores demuestran que este método funciona en tiempo real, sin embargo hasta el momento no existe evaluación de su comportamiento en términos de ajuste. En el artículo [10] se plantean modificaciones y me-

joras a este algoritmo para manejar problemas de oclusión, donde el objeto a seguir queda parcial o totalmente obstruido por otros objetos, y problemas de dispersión donde el algoritmo además de seguir al objeto de interés, sigue otros objetos con características similares. Se describe un método para detectar los casos de oclusión y se propone una modificación al algoritmo para solucionarlo. También se propone una modificación para el problema de oclusión con objetos de características similares.

Por otro lado, la evaluación de la performance de los métodos de tracking es de suma importancia, sobre todo en aquellas aplicaciones en las que la precisión de los resultados es fundamental como ocurre en el diagnóstico médico. Algunos autores han desarrollado medidas para evaluar el comportamiento de los algoritmos [5, 8, 7, 16]. En esta línea de investigación, proponemos evaluar el método de seguimiento en video presentado [10], para analizar su comportamiento.

3. Líneas de investigación y desarrollo

La idea de este trabajo es implementar medidas de evaluación del comportamiento del método de seguimiento de objetos en video que se describe a continuación, con el objetivo de medir la precisión de sus resultados y compararlo con otros algoritmos. La técnica utilizada está basada en la teoría de conjuntos de nivel. Hacemos aquí una breve descripción del mismo en el caso donde sólo hay un objeto de interés, para más información ver [17, 18, 19, 10].

Cada imagen I_k del video está compuesta por una región de interés Ω_1 , cuyo borde se quiere encontrar, y la región Ω_0 correspondiente al fondo, donde $\Omega_0 \cup \Omega_1 = I_k$ y $\Omega_0 \cap \Omega_1 = \emptyset$. Cada una de las regiones está caracterizada por los parámetros de una función de distribución $P(\mathbf{x}|\Theta_m)$ ($m = 0, 1$). Los parámetros pueden ser también los colores del objeto y del fondo. Se elige una función $\phi : R^2 \rightarrow R$ tal

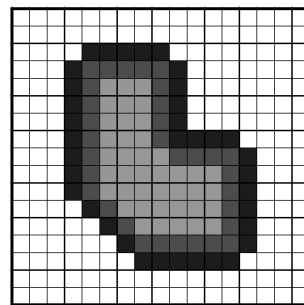


Figura 1: Objeto en una imagen, el borde interno (gris claro) y el externo (gris oscuro).

que:

$$\begin{aligned} \phi(\mathbf{x}) < 0 & \text{ si } \mathbf{x} \in \Omega_1 \\ \phi(\mathbf{x}) > 0 & \text{ si } \mathbf{x} \in \Omega_0 \end{aligned}$$

Sea $\partial\Omega_1$ el borde de la región Ω_1 que se quiere encontrar. Se definen dos listas de pixels vecinos L_{in} y L_{out} de la siguiente forma:

$$L_{in} = \{\mathbf{x} | \phi(\mathbf{x}) < 0 \text{ y } \exists \mathbf{y} \in N_4(\mathbf{x}) \text{ tal que } \phi(\mathbf{y}) > 0\}$$

$$L_{out} = \{\mathbf{x} | \phi(\mathbf{x}) > 0 \text{ y } \exists \mathbf{y} \in N_4(\mathbf{x}) \text{ tal que } \phi(\mathbf{y}) < 0\}$$

donde $N_4(\mathbf{x}) = \{\mathbf{y} / |\mathbf{x} - \mathbf{y}| = 1\}$ es el conjunto de pixels 4-vecinos de \mathbf{x} .

L_{in} y L_{out} son los bordes interno y externo del contorno $\partial\Omega_1$, respectivamente. La Figura 1 muestra un objeto en una imagen y sus bordes interno y externo, representados en diferentes tonos de gris.

En este caso la ecuación de evolución es

$$\phi_t(\mathbf{x}, t) = |\nabla\phi(\mathbf{x}, t)| (F_d + F_s) \quad (1)$$

donde

$$F_d(\mathbf{x}) = \log(P(\mathbf{x}|\Omega_1)/P(\mathbf{x}|\Omega_0)) \quad (2)$$

y

$$F_s(\mathbf{x}) = -2\lambda\kappa(\mathbf{x}) \quad (3)$$

siendo κ la curvatura.

Un pixel \mathbf{x} es interior si $\mathbf{x} \in \Omega_1$ y $\mathbf{x} \notin L_{in}$.
 Un pixel \mathbf{x} es exterior si $\mathbf{x} \in \Omega_0$ y $\mathbf{x} \notin L_{out}$.
 Entonces se define ϕ de la siguiente manera:

$$\phi(\mathbf{x}) = \begin{cases} 3 & \text{si } \mathbf{x} \text{ es un pixel exterior} \\ 1 & \text{si } \mathbf{x} \in L_{out} \\ -1 & \text{si } \mathbf{x} \in L_{in} \\ -3 & \text{si } \mathbf{x} \text{ es un pixel interior} \end{cases} \quad (4)$$

Luego se intercambian los pixels según un algoritmo de dos ciclos: en el primer ciclo se hace evolucionar el contorno siguiendo el signo de la fuerza $F_d(\mathbf{x})$. Como la fuerza $F_d(\mathbf{x})$ depende de datos de la imagen y de la caracterización del objeto, hace que el contorno inicial evolucione adaptándose al contorno del objeto. El segundo ciclo es idéntico al primero, pero utilizando la fuerza $F_s(\mathbf{x})$, que como depende de la curvatura, suaviza el contorno. En cada cuadro del video se parte de un contorno inicial, el cual evoluciona adaptándose al borde del objeto. En la primer imagen se establece arbitrariamente un contorno inicial, luego el algoritmo encuentra el borde del objeto en la imagen actual tomando como contorno inicial la curva hallada en la imagen anterior.

Cuando el objeto a seguir pasa por detrás de otro objeto y queda oculto parcial o totalmente la curva de seguimiento desaparece de la imagen o toma formas que no son correctas. El objeto de interés puede ser oculto por un objeto de color diferente o similar. En el primer caso, se produce una obstrucción y el algoritmo pierde el contorno del objeto y ya no puede seguirlo. En el segundo caso, se produce una dispersión y el algoritmo confunde el contorno del objeto. En el artículo [10] se muestran ejemplos en videos sintéticos en donde el algoritmo falla. Se presentan también las soluciones propuestas.

4. Resultados y Objetivos

El objetivo de esta línea de investigación es evaluar el comportamiento del método ante-

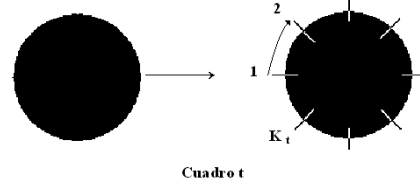


Figura 2: Rectas ortogonales al borde.

riormente descrito. Para eso comenzamos utilizando la siguiente medida δ , propuesta en [7]:

En cada cuadro t se divide el contorno en K_t segmentos de borde con rectas perpendiculares al contorno, como muestra la Figura 2.

Sean $C_{out}^i(t)$ y $C_{in}^i(t)$ los colores en un pixel exterior e interior del objeto, respectivamente, sobre la recta i . Entonces, para cada cuadro t de la secuencia y para cada segmento de recta $i = \{1, \dots, K_t\}$ se define:

$$\delta(t) = 1 - \frac{1}{K_t} \sum_{i=1}^{K_t} d(t; i)$$

donde

$$d(t; i) = \frac{\|C_{out}^i(t) - C_{in}^i(t)\|}{\sqrt{(3 \times 255^2)}}$$

Como $0 \leq \delta \leq 1$, una medición de δ cercana a cero corresponde a un seguimiento correcto, mientras que una medición cercana a 1 significa que la curva no ajusta al borde del objeto en todos los cuadros. Sin embargo esto ocurre sólo en el caso en que el contraste entre el fondo y el objeto es muy alto, como muestran los resultados de la Figura 3 y de la Tabla 1. Esta medida de evaluación de ajuste en cada frame no es adecuada para cualquier video y su eficiencia está restringida a aquellos videos de alto contraste de color entre el objeto de interés y el fondo.

En la Figura 3 se muestra una secuencia de imágenes sintéticas utilizada para probar esta medida. Se trata de un círculo desplazándose por la pantalla, testado con diferentes colores para el fondo y el objeto. En todos los casos el

ajuste de la curva resultado fue muy bueno en términos visuales, como se observa en la Figura 3. Los resultados de la medición se muestran en la Tabla 1.

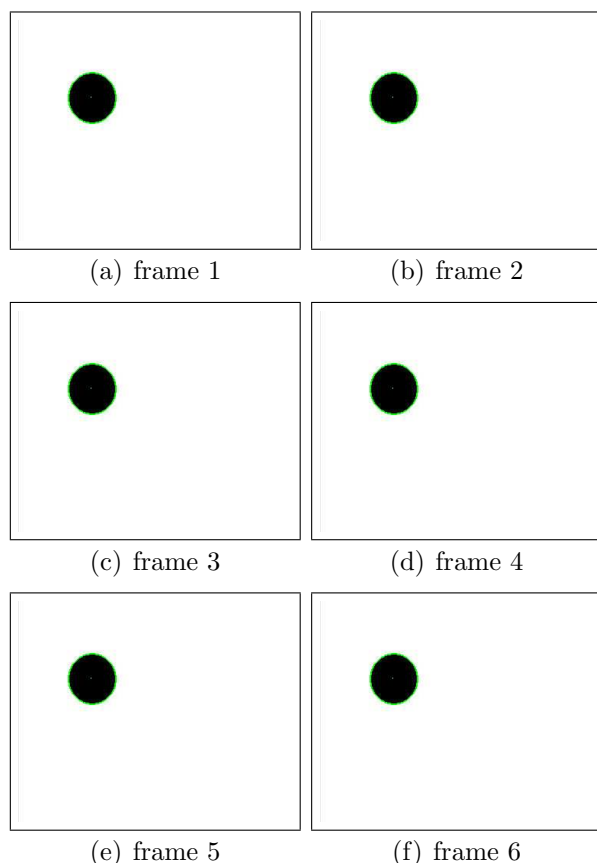


Figura 3: Resultado de aplicar el algoritmo a una secuencia de imágenes sintéticas que muestra un círculo desplazándose.

Contraste	Medida
Blanco y Negro	0.11
Rojo Oscuro y Negro	0.91
Verde Oscuro y Negro	0.92

Cuadro 1: Resultados de la medición para el video de la Figura 3.

Se observa que cuando el contraste de colores es bajo, la medida es muy grande, a pesar de que la curva ajusta muy bien y el seguimiento se realiza correctamente. Por esa razón estamos investigando otras métricas que sean más generales.

5. Formación de Recursos Humanos

Este trabajo contribuye a la formación de la Lic. Ximena Pisani.

Referencias

- [1] M.S. Allili and D. Ziou. Active contours for video object tracking using region, boundary and shape information. *Signal, Image and Video Processing*, 1(2):101–117, 2007.
- [2] Nguyen Dang Binh. A robust framework for visual object tracking. In *omputing and Communication Technologies, 2009. RIVF '09. International Conference on Digital Object*, pages 1–8, 2009.
- [3] Çigdem Eroglu Erdem. Video object segmentation and tracking using region-based statistics. *Sig. Proc.: Image Comm.*, 22(10):891–905, 2007.
- [4] T. Chan and L. Vese. Active Contours without Edges. *IEEE Trans. Image Processing*, 10(2):265–277, 2001.
- [5] D. Conte, P. Foggia, G. Percannella, and M. Vento. Performance evaluation of a people tracking system on pets2009 database. In *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, pages 119–126, 2010.
- [6] T.B. Dinh and G. Medioni. Co-training framework of generative and discriminative trackers with partial occlusion handling. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 642–649, jan. 2011.
- [7] C.E. Erdem, B. Sankur, and A.M. Tekalp. Performance measures for video object

- segmentation and tracking. *Image Processing, IEEE Transactions on*, 13(7):937–951, July 2004.
- [8] Cigdem Eroglu Erdem, A. Murat Tekalp, and Bülent Sankur. Video object tracking with feedback of performance measures. In *Circuits and Systems for Video Technology, IEEE Transactions on*, volume 13, pages 310 – 324, 2003.
- [9] O. D. Faugeras and Renaud Keriven. Variational principles, surface evolution, pdes, level set methods, and the stereo problem. *IEEE Transactions on Image Processing*, 7(3):336–344, 1998.
- [10] Juliana Gambini, Damian Rozichner, María E. Buemi, Marta Mejail, and Julio Jacobo-Berlles. Occlusion handling for object tracking using a fast level set method. In *SIBGRAPI*, pages 61–68, 2008.
- [11] Hyung-Bok Kim and Kwee-Bo Sim. A particular object tracking in an environment of multiple moving objects. In *Control Automation and Systems (ICCAS), 2010 International Conference on*, pages 1053 –1056, oct. 2010.
- [12] R. Malladi, J. A. Sethian, and B. C. Vemuri. A fast level set based algorithm for topology-independent shape modeling. *J. Math. Imaging and Vision*, 6(2/3):269–290, 1996.
- [13] Ravi Malladi, James A. Sethian, and Baba C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.
- [14] S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer, first edition, 2003.
- [15] N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: A level set approach. In *Proc. European Conf. Computer Vision*, pages 224 – 240, June 2000.
- [16] A. G. Amitha Perera, Anthony Hoogs, Chukka Srinivas, Glen Brooksby, and Wensheng Hu. Evaluation of algorithms for tracking multiple objects in video. In *AIPR*, page 35, 2006.
- [17] J. Sethian. Fast marching methods and level set methods for propagating interfaces von karman insitute lecture series, 1998.
- [18] J.A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision and Materials Sciences*. Cambridge University Press, Cambridge, 1999.
- [19] Y. Shi and W.C. Karl. Real-time tracking using level sets. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 34–41vol.2, 20-25 June 2005.
- [20] Y.G. Shi and W.C. Karl. A real-time algorithm for the approximation of level-set-based curve evolution. *IP*, 17(5):645–656, May 2008.
- [21] S. Zhu and A. Yuille. Region competition: Unifying snakes, region growing and bayes/MDL for multiband image segmentation. *IEEE Transaction on Pattern Anal. and Machine Intelligence*, 18(9):884–900, 1996.